

KONE AVUKSI SISÄLLÖNKUVAILUUN? YLEN AUTOMAATTISEN SISÄLLÖNKUVAILUN HANKE

Matilda Honkajuuri

Tampereen yliopisto
Viestintätieteiden tiedekunta
Informaatiotutkimus ja interaktiivinen media
Pro gradu -tutkielma
Huhtikuu 2018

TAMPEREEN YLIOPISTO, Viestintätieteiden tiedekunta

Informaatiotutkimus ja interaktiivinen media

HONKAJUURI, MATILDA: Kone avuksi sisällönkuvailuun? Ylen automaattisen sisällönkuvailun hanke

Pro gradu -tutkielma, 72 s., 2 liitettä.

Huhtikuu 2018

TIIVISTELMÄ

Tässä tutkielmassa käsitellään Ylen automaattisen sisällönkuvailun hanketta vuodelta 2016. Ylellä on mittavat multimedialliset kokoelmat, joiden omatuotantoinen osa on läikin perustuen arkistoitava pysyvästi riittävin metatiedoin. Jotta arkistointi takaisi multimediallisten aineistojen mahdollisimman laajan jälleenkäyttöarvon, aineistoille tehdään manuaalisesti sisällönkuvailu. Sisällönkuvailu tehdään luonnollisella kielellä kokonaisia lauseita käyttäen. Ihmistyönä tehtävä sisällönkuvailu on hidasta ja kallista. Osittain tämän vuoksi Ylessä aloitettiin hanke, jonka tarkoituksena oli koettaa tehostaisiko ja helpottaisiko automaattinen sisällönkuvailu sisällönkuvailijoiden työtä.

Automaattisen sisällönkuvailun ohjelman Ylen hankkeeseen tuotti Valossa Labs Oy. Valossa Labs Oyn testaukseen tuottama ohjelma oli vasta kehityksessä oleva sovellus, eikä valmis automaattisen sisällönkuvailun ohjelma. Hankkeen aikana ohjelmaa testattiin Ylen puolelta, käyttäen testauksessa Ylen aineistoja. Tutkielma keskittyy sisällönkuvailun ammattilaisten käyttäjäkokemuksiin hankkeen ajalta.

Ylen automaattisen sisällönkuvailun hanketta on tässä tutkielmassa tarkasteltu käyttäjälähtöisten menetelmien avulla. Tutkielmassa selvitetään, miten sisällönkuvailijat kokevat sisällönkuvailun, joka tehdään manuaalisesti ja miten he kokivat automaattisen sisällönkuvailun mahdollisena osana työtään.

Tulosten perusteella voidaan sanoa, että automaattinen sisällönkuvailu ei vielä tuota Ylen tarpeisiin riittävän laadukasta sisällönkuvailua. Automaattinen sisällönkuvailu ei ole sillä tasolla, että se tehostaisi sisällönkuvailijoiden työtä. Kuitenkin voidaan esittää oletus siitä, että automaattinen sisällönkuvailu saattaisi tuottaa sellaista metadataa, jolle varsinkin julkaisu- tai tiedonhakutilanteissa olisi tarvetta. Automaattinen sisällönkuvailu ei kykene tuottamaan semanttista kuvailua, joka riittäisi Ylen tarpeisiin, mutta se voisi tuottaa lisäarvoa tiedonhakuun. Johtopäätös on se, että automaattisen sisällönkuvailun ja ihmislähtöisen sisällönkuvailun yhdistäminen tuottaisi monipuolista ja semanttisesti rikasta sisällönkuvailua.

Avainsanat: automaattinen sisällönkuvailu, multimedialliset aineistot, videoaineistot, arkistointi, käyttäjälähtöinen tutkimus

Sisällysluettelo

1	JOHDANTO	1
2	KIRJALLISUUSKATSAUS	3
	2.1 Kuva- ja videoaineistojen tiedonhaun tutkimuksia.....	3
	2.2 Multimedialliset aineistot sisältävät useaa mediaa	5
3	AUTOMAATTINEN SISÄLLÖNKUVAILU	7
	3.1 Automaattinen sisällönkuvailu ja kuvantunnistus	7
	3.2 Automaattisen sisällönkuvaailun tarve Ylellä	9
	3.3 Automaattisen sisällönkuvaailun testaus	11
4	METODIT	13
	4.1 Tutkimuskysymykset	13
	4.2 Haastatteluaaineiston keruu	14
	4.3 Affinity diagram -metodi	17
	4.4 Empiirisen aineiston analyysi	20
5	SISÄLLÖNKUVAILUN KÄYTÄNNÖT YLELLÄ	24
	5.1 Sisältöjä pitäisi voida kuvailla ilman formaalia koulutusta	24
	5.2 Sisällönkuvaailijana haluan tehdä työni tehokkaasti	27
	5.3 Aineistot eivät ole homogeenisiä	30
	5.4 Hyvin tehty sisällönkuvailu mahdollistaa jälleenkäytön	32
	5.5 Videoaineistojen sisällönkuvailu ja aikadimensio	36
	5.6 Sisällönkuvailu vaatii tietojen yhdistämistä	42
6	MIKSI KONE EI RIITÄ SISÄLLÖNKUVAILUSSA.....	48
	6.1 Ihminen kertoo sen mitä kuvassa ei näy	48
	6.2 Automaattisen sisällönkuvaailun oppimiskyky on rajallinen	51
7	IHMISEN JA KONEEN YHTEISTYÖ	57
	7.1 Kone ei vielä korvaa ihmistä	57
	7.2 Automaattinen sisällönkuvailu voisi tasalaatuistaa sisällönkuvaailuja	61
	7.3 Vielä ei olla valmiita automaattisen sisällönkuvaailun tuomiin muutoksiin	65
8	YHTEENVETO	68
	8.1 Johtopäätökset.....	68
	8.2 Jatkotutkimus	71
	LÄHTEET	72
	LIITTEET	

1 JOHDANTO

Yleisradiolla on lakisääteinen velvollisuus tallentaa kaikki omatuotantoiset videoaineistonsa.

”Tämän lain tarkoituksena on Suomessa yleisön saataville saatettujen kansallisen kulttuurin aineistojen säilyttäminen tuleville sukupolville ja saattaminen tutkijoiden ja muiden tarvitsijoiden käyttöön.” (Laki kulttuuriaineistojen tallettamisesta ja säilyttämisestä, 28.12.2007/1433).

Laki elokuvien arkistoinnista on ollut voimassa vuodesta 1984 ja siitä lähtien kaikki Ylen omatuotantoiset ja yhteistuotannolliset ohjelmat on arkistoitu. Pelkkä laki ei kuitenkaan motivoi arkistointiin, vaan taustalla on myös tarve hyödyntää olemassa olevaa aineistoa uusien ohjelmien tekemisessä. Ennen lain voimaan astumista Ylellä on arkistoitu omatuotantoisia videoaineistoja vuodesta 1957 lähtien. Yle on myös hankkinut aineistoa elokuvayhtiöiltä, joista vanhimmat ovat vuodelta 1906. (Yle 2017).

Nykyisin Ylellä on käytössä medianhallintajärjestelmä, johon on tallennettu kaikki videoaineisto ja niitä kuvaileva metatieto. Vuodesta 2009 lähtien kaikki aineisto on tallennettu digitaalisena ja vanhempaa aineistoa digitoidaan jatkuvasti. Ylessä arkistoitavat multimeditalliset aineistot ovat lähinnä tv-aineistoja, mutta kokoelma sisältää myös esimerkiksi työmateriaalia sekä internetissä julkaistuja klippejä. (Kajo 2018).

Videoarkistoinnin pitkän historian ja lain pakottavuuden vuoksi Ylen videoaineistojen arkistot ovat hyvin laajat. Jotta aineistot olisivat löydettävissä, niille tulee tehdä riittävän hyvä sisällönkuvailu, jolla mahdollistetaan aineistojen jälleenkäyttö. Ylen työntekijät kuvailevat videoaineistot ja niiden kuvasisällön pääasiassa manuaalisesti. Tähän työhön sisältyy paljon rutiininomaisia vaiheita, joita osa työntekijöistä on pitänyt aikaa vievinä ja jossain määrin jopa turhauttavina. Töiden tekemisen helpottamiseksi Ylessä ja sisällönkuvailun laadun tasaamiseksi Yle aloitti vuonna 2016 hakeen, jonka tarkoituksena oli tutkia voisiko organisaatiolle olla hyötyä automaattisesta sisällönkuvailuohjelmasta. Hankkeessa testattiin Valossa-yrityksen kehittelyvaiheessa olevaa sovellusta.

Tämä tutkimus kohdentuu Ylen automaattisen sisällönkuvailun hankkeeseen. Tutkimuksesta selviää, miten sisällönkuvailun parissa työskentelevät henkilöt kokivat automaattisen sisällönkuvailun verrattuna nykyiseen sisällönkuvailuun. Tutkielman keskeisiä kysymyksiä ovat, millaista sisällönkuvailua automaattisen sisällönkuvailun ohjelma tuotti ja

miten automaattisen sisällönkuvailun tuottamaa sisällönkuvailua voitaisiin hyödyntää tulevaisuudessa. Tutkimus on tehty laadullisena haastattelututkimuksena, käyttäjäkeskeisiä menetelmiä soveltaen.

Tutkielma rakentuu seuraavasti. Luvussa 2 tarkastellaan tutkimusta taustoittavaa kirjallisuutta. Luku 3 keskittyy automaattisen sisällönkuvailun piirteisiin. Tutkielmassa käytetyt menetöt käsitellään luvussa 4. Aineiston analyysi alkaa luvusta 5, jossa käsitellään Ylen sisällönkuvailun käytäntöjä. Luvussa 6 pohditaan miksi automaattinen sisällönkuvailu ei riitä kattamaan sisällönkuvailun tarvetta. Manuaalisen ja automaattisen sisällönkuvailun yhteistyön mahdollisuuksia pohditaan luvussa 7. Luvussa 8 kerrotaan tutkielman johtopäätökset sekä esitellään mahdollisia jatkotutkimusaiheita.

2 KIRJALLISUUSKATSAUS

Tässä luvussa käydään läpi aiempaa tutkimusta videotiedonhausta, kuten multimediallisten aineistojen semanttisesta sisältöanalyysistä. Tämän jälkeen selvitetään, mitä multimedialliset aineistot ovat sekä määritellään tarkemmin, mitä multimediallisilla aineistoilla tarkoitetaan.

2.1 Kuva- ja videoaineistojen tiedonhaun tutkimuksia

Tiedonhaun tutkimuksen keskeisiin kysymyksiin kuuluu, miten tiedonhakujärjestelmistä saadaan haettua relevanttia informaatiota (Vakkari, 1999). Tiedonhakujärjestelmät ovat tietokoneohjelmistoja, jotka säilyttävät ja hallinnoivat informaatiota kuten dokumentteja, mutta myös multimediallisia aineistoja. Tiedonhakujärjestelmät auttavat käyttäjää löytämään hänelle relevanttia informaatiota. Tarkoituksena on täyttää käyttäjän tiedontarve. Täydellinen tiedonhakujärjestelmä hakee käyttäjälle vain relevantteja aineistoja käyttäjän muodostaman kyselyn avulla. (Ridley, Goker & Davies, 2009).

Kuvatiedostoihin kohdistuva tiedonhaku on ollut tutkimuksen kohteena jo yli 50 vuotta. Pitkästä ja merkittävästä kuvatiedostojen tiedonhaun tutkimuksen historiasta huolimatta, verrattain vähän tutkijoita on ollut konenäön analyysin eturintamassa. Huomionarvoista on todeta, että suurin osa tutkimuksesta, joka kohdistuu kuvatiedostojen ja varsinkin videoaineistojen tiedonhakuun sijoittuu 1990-luvulle.

1990-luvulla tapahtunut teknologinen kehitys toi kuvatiedonhaun tietojenkäsittelytieteiden kiinnostuksen kohteeksi. Nykyisin kuvien ja videoiden määrä on suuressa kasvussa. Tähän on vaikuttanut internetin kasvanut suosio, mobiilialustojen kehittyminen sekä internetissä tapahtuva sosiaalinen verkostoituminen. Kuvien ja videomateriaalien käyttäminen on levinnyt laajasti esimerkiksi Googlen kuvahaun, videomateriaalien jakamiseen tarkoitetun sivuston YouTuben, televisiotuotantoyhtiöiden suoratoistopalveluiden sekä henkilökohtaisten valokuva-arkistojen laajan kasvun myötä. Tutkimus on siirtynyt syntaktisesta sisällöstä semanttisen sisällön tutkimiseen. (Enser 2008, 531-533). Kuitenkin tutkimus on toistaiseksi verrattain vähäistä.

1990-luvulla moni tutkija keskittyi videoaineistojen tiedonhaun ongelmiin, etenkin järjestelmien mahdollisuuteen tuottaa sisällönkuvailua korkeammalla tasolla. Aiheita käsitelivät tutkimuksissaan esimerkiksi: Brunelli, Mich ja Modena (1999). He kirjoittivat vuonna 1999 automaattisen videoanalyysin olevan todistettavasti hyödyllistä multimediallisten aineistojen hallintaan. Samalla he kuitenkin totesivat, että tarvitaan paljon työtä kaikkien videoanalyysin tutkimuksen saroilla, jotta multimediallisten aineistojen hakuun löydetään samanlaista helppoutta kuin mitä tekstiedonhaussa on. (Brunelli et al. 1999, 106). Samana vuonna Colombo, Del Bimbo ja Pala (1999) julkaisivat artikkelin ”Semantics in visual information retrieval”, jossa todetaan visuaalisten tiedonhakujärjestelmien astuneen uuteen aikakauteen. Artikkelin päälöydös oli, että sääntöjen määrittäminen visuaalisesti kuvatulla semantiikalla on vaikeaa, varsinkin kun puhutaan tavallisista kuvista. (Colombo et al. 1999).

Vuosituhanneen vaihteen jälkeen automaattisen videoanalyysin tutkimus kääntyi laskuun. Vaikka tutkimus on jäänyt niukaksi, on aihe kuitenkin kiinnostanut muutamia tutkijoita. Aiheellisesti tutkimukset kääntyivät enemmänkin semanttisen kuilun ongelmien ratkaisuun ja tutkimiseen. Videoaineistoja ja semantiikkaa on tutkinut muun muassa Smith (2001). Hän on osoittanut, että videoaineistojen visuaalinen informaatio voidaan jakaa kolmeen eri tasoon: semantiikka, rakenne ja piirteet. Myös Hare, Lewis, Enser sekä Sandom (2006) ovat tutkineet semanttista kuilua videoaineistojen tiedonhaussa. Heidän mukaansa semanttinen kuilu voidaan jakaa kahteen pääluokkaan: kuiluun, joka sijaitsee kuvailun ja videoaineistosta löytyvien objektien välillä, sekä kuiluun, joka sijaitsee nimitöiden objektien ja täydellisen semantiikan välillä. (Hare et al., 2006).

Enser (1999) on tutkinut myös käsitteellisen ja sisältöperustaisen tiedonhaun eroja visuaalisten aineistojen kohdalla. Hän pyrki selittämään, miksi visuaalisessa tiedonhaussa halutaan edelleen painottaa sisältöperustaista, eikä käsitteellisyyteen perustuvaa tiedonhakua. Tutkimuksen tuloksena esitellään ajatus siitä, että hyvä yhdistelmä, sekä käsitteellistä että sisältöperusteista tiedonhakujärjestelmää tuottaisi tiedonhakijan kannalta parhaan lopputuloksen. (Enser, 1999)

Konenäköön perustuvaa sisällönkuvailua ovat tutkineet mm. Liu, Zhang, Lu ja Ma (2007). Heidän mukaansa semanttinen kuilu johtuu siitä, että ihmisellä on tapana käyttää korkeatasoisia konsepteja. Vastaavasti automaattinen konenäköön perustuva tekniikka luo matalan tason sisältöjä, kuten värejä, tekstiä ja muotoja.

Ihmisen subjektiivista suhtautumista videoaineistojen kuvailemiseen ovat tutkineet Gotoh ja Khan. Heidän mukaansa siis kaksi ihmiskuvailijaa tuottavat kaksi hyvin erilaista kuvausta samasta videoaineistosta. Tätä kuilua voidaan heidän mukaansa kaventaa käyttämällä avainsanoja. Gotohin ja Khanin (2016, 244-246) mukaan avainsanojen käyttäminen on hyödyllistä, kun halutaan tunnistaa objekteja tai toimintoja videomateriaalista. Avainsanojen käyttämisen lisäksi tutkimuksissa on käsitelty myös tägeillä vaikutusta videoaineistojen sisällönkuvailuun. Käyttäjien lisäämien tágien merkitystä, laatua sekä arvoa kuvatiedonhaun semantiikkaan on tutkinut mm. Jörgensen, Stivilia ja Wu (2014). Tutkimuksen mukaan objekteja, ihmisiä sekä paikkoja kuvaillaan eniten tágien avulla. Jörgensenin, Stivilian ja Wun mukaan tágellä pystytään kuvailemaan kohteita informatiivisesti kuvasta, mutta käyttäjät arvostavat myös kontrolloituja sanastoja, koska ne herättävät luottamusta.

Zhou ja Huand (2002) ovat tutkineet avainsanojen käyttämistä kuvatiedonhaussa. Heidän mukaansa avainsanaan perustuva tiedonhaku tarvitsee taustalle tesauruksen, muuten avainsanahaku jää liian suppeaksi. Tehokas tiedonhakuprosessi on riippuvainen kuvailusta, jossa käyttäjä toimii luokittelun sekä toisten käyttäjien kanssa johdonmukaisella tavalla. Avainsanojen käyttämisessä ilmenee ongelmia, kun käyttäjät kuvailevat kuvaa samankaltaisilla sanoilla, jotka kuitenkin eivät ole samoja. Ongelmia syntyy myös, kun käyttäjä ei käytä juuri kuvailussa käytettyä sanaa kuvaamaan haettua objektia, vaan esimerkiksi sanan synonyymia. Edellä mainittuihin ongelmiin auttavat tutkimuksen mukaan esimerkiksi tesaurukset, koska ne määrittelevät kuvailussa käytettävät termit.

2.2 Multimedialliset aineistot sisältävät useaa mediaa

Multimedialliset aineistot ovat aineistoja, joissa määritelmän mukaan yhdistyvät vähintään kaksi seuraavista: teksti, grafiikka, animaatio, liikkuva kuva, liikkumaton kuva tai ääni (Rautiainen, 2006, 28). Koska tässä tutkielmassa käsitellään enimmäkseen aineistoja, jotka sisältävät ainakin liikkuvaa kuvaa ja ääntä, niitä voidaan kutsua myös termillä videoaineistot. Videoaineistot ovat Tannin (2003, 8) määritelmän mukaan multimodaalisia ja sekventiaalisia, yksittäisistä kuvista muodostuvia kuvavirtoja. Tannin tutkimuksessa multimodaalisilla aineistoilla tarkoitetaan aineistoja, jotka sisältävät useaa mediaa, esimerkiksi kuvaa, tekstiä ja ääntä. Tanni viittaa Ruin ja kumppaneiden tutkimukseen,

jossa sekventiaalisuudella tarkoitetaan sitä, että kaksi tai useampi samanaikaisesti esiintyvää prosessia on käytännössä esitettävä vuorotellen. Kuvavirrat koostuvat useiden yksittäisten kuvien sarjasta, jotka on esitetty niin tiheästi, ettei ihmissilmä niitä välttämättä pysty erottamaan toisistaan (Tanni, 2003, 9). Kuitenkin kuvavirtoja voidaan katkoa esimerkiksi kohtauksiksi tai muiksi ohjelman osuuksiksi. Tannin (2003, 9) mukaan tämä kuvavirtojen katkonaisuus on videon mahdollisesti keskeisin rakenteellinen ominaispiirre. Yle hyödyntää ohjelmien katkonaista rakenteellisuutta jakamalla ohjelmat segmentteihin. Uutiset pilkotaan osuuksiin, joista jokainen on yksittäinen uutinen. Näitä katkonaisuudesta osiksi pilkottuja palasia kutsutaan segmenteiksi ja rajoja segmenttirajoiksi. Segmentit koostuvat kuvavirroista ja kuvavirrat luovat videolle myös ajallisen dimension.

Video on siis ainestoa, joka muodostuu useasta eri median muodosta, joka taas koostuu useasta hyvin nopeasti peräkkäin esitetyistä kuvasta ja jolla on tämän vuoksi myös ajallinen dimensio. Koska tässä tutkimuksessa keskitytään multimediallisista aineistoista nimenomaan videoaineistoihin, on otettava huomioon kuvien lisäksi myös audio. Audion voidaan ajatella lisäävän videoaineistoon yhden kerroksen. Rautiaisen (2006, 31) mukaan audioraita ei välttämättä aina kulje aivan yhtäaikaisesti videon kanssa, vaan se kuljettaa tarinaa eri tasolla, esimerkkinä Rautiainen mainitsee narratiivisen kerronnan dokumenttielokuvissa. Audion voidaan siis olettaa täydentävän kuvallista kerrontaa, vaikka kuvallinen kerronta ja ääni, eivät välttämättä kerro aina samaa tarinaa synkronoidusti.

Multimediallisille aineistoille on tyypillistä myös ainestoista tulkittavissa oleva semantiikka. Smith (2001, 971) esittelee artikkelissaan ”Quantitative assessment of image retrieval effectiveness”, kuinka visuaalinen informaatio voidaan jakaa kolmeen eri tasoon. Nämä tasot ovat semantiikka, rakenne ja piirteet. Semantiikka on tämän määritelmän mukaan kontekstisidonnaista, objektilla (”jalkapalloilija”) on suhde (”potkaista”) toisen objektin (”pallo”) kanssa.

3 AUTOMAATTINEN SISÄLLÖNKUVAILU

Tässä luvussa luonnehditaan aluksi automaattisen sisällönkuvailun piirteitä ja selostetaan, mitä automaattisen sisällönkuvailun ohjelma tekee. Tämän jälkeen tarkastellaan automaattisen sisällönkuvailun tarpeita Ylellä ja kuvataan testissä hyödynnetyn automaattisen sisällönkuvailun ohjelman ominaisuuksia.

3.1 Automaattinen sisällönkuvailu ja kuvantunnistus

Automaattisessa sisällön analyysissä nykyisin hyödynnettävät kuvantunnistuksen välineet tukevat monissa tapauksissa onnistuneesti semanttisten konseptien kuten ihmisten, rakennusten, vertailevasti luonnon näkymien sekä ihmisen rankentamien näkymien, tyydyttävällä tasolla olevaa tunnistamista (Dalakleidi, et al, 2011 s.196). Automaattiset sisällönkuvailun ohjelmat pystyvät johonkin tasoon asti tunnistamaan kuvassa olevia objekteja, mutta sisällönkuvailun ohjelmat eivät ole vielä sillä tasolla, että ne ymmärtäisivät kuvailtavana olevien objektien semanttisen tason. Tiedonhakija voisi esimerkiksi haluta videokuvaa kahdesta koirasta, jotka leikkivät pallolla. Automaattisen sisällönkuvailun ohjelma pystyy tunnistamaan kuvasta kolme tarvittavaa objektia: koiran, toisen koiran sekä pallon. Automaattinen sisällönkuvailun ohjelma siis kyllä tunnistaa pallolla leikkivät koirat, mutta ei pysty tuottamaan sisältöä, että koirat nimenomaan leikkivät pallolla kuvassa. Ohjelma pystyy tunnistamaan kuvasta vain kolme irrallista objektia, ei sitä mikä on objektien keskinäinen suhde. Kuvan tunnistus ei siis pysty vielä tuottamaan sisältöjä kuvien merkityksestä, eli semantiikasta.

Kuvantunnistuksen on vaikea tunnistaa merkityssuhteita, koska kuvan attribuuttina merkitys on hyvin erityyppinen verrattuna muihin kuvan ominaisuuksiin. Merkitys on huonosti määriteltävissä, toisin kuin esimerkiksi väri. (Enser, 1999). Kuvassa on useita attribuutteja, kuten edellä mainittu väri. Kuvan attribuutit luovat kuvalle merkityksiä, riippuen siitä, miten ne ovat suhteessa keskenään. Merkityksiä voi kuvassa olla usealla eri tasolla. Enser ja Sandomin (2002) lainaavat taidehistorioitsija Panofskyä vuodelta 1962, jossa Panofsky jakaa kuvan olemuksen kolmelle tasolle, jolla sen sisältämää aihetta tai käsitteitä voidaan analysoida. Enserin ja Sandomin mukaan nämä Panofskyn esittämät kuvan kolme tasoa ovat relevantteja, myös kun puhutaan liikkuvasta kuvasta, eli videoaineistoista. Tasot ovat seuraavat:

- 1) Esi-ikonografinen taso. Viittaa siihen mitä kuva sisältää yleisellä tasolla, esimerkiksi nainen, vauva tai rakennus.
 - 2) Ikonografinen taso. Viittaa spesifeihin tunnistuksiin, esimerkiksi paikkojen ja henkilöiden nimeämiseen, Madonna ja lapsi tai Taj Mahal.
 - 3) Ikonologinen taso. Viittaa kuvan symboliikkaan, abstrakteihin konsepteihin kuten toivo tai rakkaus.
- (Enser, Sandom, 2002)

Multimediallisen kuvan kuvailuun pitää siis liittää, ei vain nimikoituja asioita, vaan myös syvempää ymmärrystä siitä, mitä kuvassa on, mikä on kuvassa olevien kohteiden keskinäinen suhde, suhde maailman tilaan sekä muuhun kontekstiin. Hare et al. (2006) kirjoittavat, että me voimme nimetä objekteja kuvasta, mutta niiden merkitys tai koko kuvan merkitys jää epäselväksi.

Toisaalta joissain tapauksissa on mahdollista, että kuvan koko semantiikka ja merkitys voidaan saavuttaa vain listaamalla kuvasta tunnistettuja objekteja. Hare et al. (2006) kirjoittavat, että joissakin tapauksissa vaadittu semantiikka kyselyissä voidaan ilmaista vain listaamalla objektien nimiä, mutta usein ilmaistut semantiikat kyselyissä ovat korkeammalla tasolla, kuin vain listatuissa nimissä. Kyse on siis siitä, että useimmiten kuvailuun tarvitaan myös muuta, kuin konenäöllisesti tunnistettuja objekteja, jotta tiedonhakijan semanttinen vaatimus saataisiin täytettyä.

Dalakeidi et al. (2011, 199) tiivistävät asian hyvin toteamalla, että kuvan ja videon kuvailussa on hyvä liittää yhteen kuvailun eri ulottuvuuksia, eikä vain käsiteltävän aineiston sisältöä, mikä on yleinen käytäntö tekstiaineistojen kohdalla. Rautiainen (2006, 52) huomauttaa väitöskirjassaan ”Content-based search and browsing in semantic multimedia retrieval”, että on vaikeaa löytää kaiken kattavaa visuaalista esimerkkiä vaikkapa sellaisille aiheille, kuten kuvaus ympäristön pilaantumisesta. Tämän vuoksi videon sisällön analyysin tulisi olla semanttisesti merkitsevämpää, symbolista esitystä, joka vähentää semanttisen kuilun olemassaoloa raakadata ja kielellisen tulkinnan välillä. Raakadata tässä yhteydessä tarkoittaa esimerkiksi listaa tunnistetuista objekteista, jotka eivät ole liitoksissa toisiinsa tai ohjelman segmenttien kestoajaksi. Raakadata on siis metatietoa videosta. Tarvittaisiin siis konenäkö, joka tunnistaisi kuvasta osia ja osaisi liittää niiden merkityssuhteet yhteen näiden vangiten kaiken oleellisen, jota tarvitaan kuvan semanttiseen kuvailuun.

Dalakleidi et al. (2011, 20) korostavat, että tiedonvälittäjän tarkoitus ei ole kommunikoida vain monimutkaisten multimodaalisten informaatiokanavien kautta. Lisäksi pitää kommunikoida myös epäsuorien sivumerkitysten, kertomusten ja puheiden kautta, jotka luovat uuden merkityksen. Jotta voidaan luoda sovelluksia ja palveluita, jotka ymmärtävät semantiikkaa, sekä multimediallisen sisällön että kontekstin pitää olla selkeää. Varsinkin puhuttaessa multimediallisista aineistoista, kuten videoaineistosta, jotain oleellista jää puuttumaan jos kuvailussa keskitytään vain asioiden tunnistamiseen. Näin ollen konteksti ja kaikki muu mediallyinen sisältö, kuten vaikkapa äänet jäävät kokonaan kuvailusta ulkopuolelle. Voidaan jopa kyseenalaistaa, onko kyseessä silloin enää multimediallinen kuvailu, jos muita medioita kuin kuva, ei oteta sisällönkuvailussa huomioon. Näin vajavaisella kuvailulla on vaikeaa saada aikaan täysin semanttista kuvailua. Tällöin aineiston ja käyttäjän välille syntyy semanttinen kuilu. Tarpeeksi hyvä kuvailu on välttämätöntä aineiston löydettävyyden takaamiseksi. Hyvä kuvailu myös helpottaa sisällön semanttisen puolen ja automaattisen kuvailun erottamista toisistaan (Dalakleidi, et.al, 201, 18).

Videoaineistojen kuvailussa on tarpeen kiinnittää huomiota myös niiden tekstillisiin osuuksiin. Merkkijonomuodossa esitettyjä metatietoja ovat esimerkiksi tekstitys, otsikko, genre, henkilöiden nimet (hahmot, näyttelijät) sekä joissain tapauksissa synopsis. Tärkein merkkijonomuotoinen aineisto tässä tutkimuksessa käsiteltylle automaattisen sisällönkuvailun ohjelmalle oli tekstitystiedostot. Tekstitystiedostojen avulla ohjelma pyrki päättämään ohjelman aihetta sekä muuta sisältöä. Ylen automaattisen sisällönkuvailun hankkeessa käytetty ohjelma käytti sisällönkuvailussaan tekstitystiedoston lisäksi myös plansitekstien tunnistusta, joka voidaan tässä yhteydessä lukea merkkijonomuotoisen sisällönkuvailun analyysiin. Käytännössä ohjelma tunnisti esimerkiksi uutisissa haastateltavina olleiden henkilöiden ruutuun tulevat tiedot, kuten nimen tai tittelin.

3.2 Automaattisen sisällönkuvailun tarve Ylellä

Erityyppisiä aineistoja voidaan luokitella kahdesta syystä (Chowdhury & Chowdhury 2007, 8). Ensinnäkin siksi, että halutaan tunnistaa aineistoja, joilla on yhteneväisiä ominaisuuksia. Toiseksi halutaan erottaa aineistot, joilla on erityyppisiä ominaisuuksia. Näin voidaan paikantaa haetun aineiston koko aineistomassan joukosta. Aineistojen luokittelu

auttaa aineistojen organisoinnissa, näyttää aineistojen väliset suhteet ja auttaa muodostamaan mentaalisen kartan siitä millaisia aineistoja kokoelmaan kuuluu.

Multimediallisten aineistojen määrä on kasvanut paljon. Esimerkiksi Yle Areenassa julkaistaan vuosittain noin 15 000 tuntia videosisältöä (Viljanen & Mattila, 2017). Tämä valtava aineistomäärän kasvu on johtanut siihen, että aineistoja halutaan prosessoida ja hallita sisältöjä semanttisesti rikkaalla tavalla (Dalakleidi, et.al., 2011, 18). Ylessä sisällönkuvailun ja metadatan lisäämiselle on kaksi hiukan erilaista tarvetta. Ensinnäkin tarve on kuluttajien parempi palveleminen. Toisena tarpeena on sisäisen ohjelmatuotannon tukeminen, eli aineistojen parempi löydettävyyden tietokannoista sekä aineistojen hyödyntäminen aineiston uudelleen käytön näkökulmasta. Ihmisen tekemä sisällönkuvailun työ on erittäin aikaa vievää, eikä kaikkia aineistoja ehditä kuvailemaan erityistä tarkkuutta noudattaen. Viljanen ja Mattila huomauttavat, että metatiedon tuottaminen käsityönä on hidas ja työlästä, koska siinä joudutaan keskittymään ydinasioihin, eikä löydettävyyden maksimointiin (Viljanen & Mattila, 2017). Työläys johtaa myös siihen, että Ylen videoaineistojen arkistossa joudutaan priorisoimaan ohjelmakohtaisesti, miten tarkkaan kuvailu suoritetaan. Aineistojen suuri määrä ja sisällönkuvailun hitaus ovat johtaneet väistämättä resurssipulaan.

Ylen verkkovälinekehitystiimin tuottajan Pia Virtasen mukaan resurssit eivät enää riitä videoaineistojen sisällönkuvailuun ihmistyönä tehtynä. Kun aineistoja ei resurssipulan vuoksi kuvailla riittävästi aineistojen arvo laskee, jos niitä ei ole kuvailtu tekstimuotoisella sisältömetatiedolla. Ilman sisältömetatietoa emme voi tietää, minkälainen on esimerkiksi multimediallisen aineiston sisältö, eikä aineisto ole ilman sisältöanalyysiä löydettävissä arkistosta. Virtanen (2016) kertoo Ylen toivoneen saavansa automaattisen sisällönkuvailun avulla enemmän parempaa ja tasalaatuisempaa kuvailua multimediallista aineistoista. Kaiken takana on kuitenkin kuluttajien parempi palveleminen. Tehokkaammin tuotetun ja laadukkaamman sisältöanalyysin avulla Yle voi tarjota kuluttajilleen paremmin sisältöjä käytettäväksi.

Ylen kehityspäällikkö Saarikosken mukaan Yle haluaa toimia suunnannäyttäjänä tekoälyn hyödyntämisessä. Suomen kieli on harvinainen maailmanlaajuisesti katsottuna, joten Suomi maana ei ole houkutteleva suurille teknologiayrityksille, jotka keskittyvät tekoälyn kehittämiseen. Tämän vuoksi Yle kokee velvollisuudekseen huolehtia osaltaan siitä, että

Suomessa media-aloilla on ajantasaiset ja maailmanlaajuisesti kilpailukykyiset työvälineet. (Saarikoski, 2018)

Ylellä uskottiin ajan olevan kypsä automaattista sisällönkuvailua varten. Tekniikoiden uskottiin olevan sillä tasolla, että automaattinen sisällönkuvailu pystyy tehostamaan sisällönkuvailijoiden työtä. (Virtanen, 2016). Tämän vuoksi Yle lähti kokeilemaan saataisiinko sisällönkuvailun ongelmiin apua automaattisesta sisällönkuvailusta.

3.3 Automaattisen sisällönkuvailun testaus

Automaattinen sisällönkuvailu tässä tutkielmassa tarkoittaa kuvailua, jonka ohjelmisto tekee videoaineistolle. Tapausesimerkissä Ylen PoC-testissä, niin sanotussa demovaiheen testissä (Proof of Concept), olleen Valossa Labs Oy:n kehittämän sisältöraportti-työkalu käytti seuraavia tunnistusmenetelmiä:

- tekstitystiedoston analyysi ja asiasanoitus sen pohjalta
- kuvantunnistus yleisesti kuvassa esiintyvien asioiden pohjalta.
- kuvantunnistus henkilöiden pohjalta
- kuvapinnan ocr-planssitekstien osalta
- kuvan automaattinen jakaminen otoksiin

(Saarikoski, 2016)

On tarpeellista huomata, että testattavana ollut sisällönkuvailun työkalu oli kehitystyön alkuvaiheessa oleva versio. Ylen tarkoituksena oli kokeilla, miten automaattiset menetelmät toimivat heidän toimintaympäristössään ja työnkulussaan. Näiden menetelmien testaamista varten Valossa Labs Oy rakensi automaattisen sisällönkuvailun pilotti-/testisovelluksen, jota ei ollut tarkoitus sellaisenaan viedä tuotantoon. Valossa Labs Oy:n tarkoituksena oli saada loppukäyttäjiltä käyttökokemuspalautea, jotta sisällönkuvailun ohjelmaa voitaisiin kehittää loppukäyttäjät huomioiden.

Testissä ollut automaattisen sisällönkuvailun ohjelma tuotti siis mm. asiasanoja, objekteja, nimiä sekä titteleitä. Näiden lisäksi ohjelman taustalla pyöri tekoälypalvelu, jonka vuoksi ohjelma pystyi oppimaan saadun käyttäjäpalautteen perusteella siitä, kuinka hyvin tunnistus onnistui.

Ylen automaattisen sisällönkuvailun testaus aloitettiin lokakuussa 2016 ja testaus jatkui saman vuoden joulukuun alkuun asti. Testaukseen osallistui Ylen työntekijöitä kaikkiaan kuusi, neljä tuotantokoordinaattoria, yksi informaatikko ja Ylen kehityspäällikkö. Automaattisen sisällönkuvailun ohjelman toimitti Valossa Labs Oy. Ylen asiakasrajapinnassa Valossa Labs Oy:n työntekijöitä työskenteli kolme henkilöä: yrityksen toimitusjohtaja sekä kaksi sovelluskehittäjää. Testauksessa käytettiin Yle Arkiston videoaineistoja. Automaattisen sisällönkuvailun testiohjelmiksi valikoituivat A-studio, A-studio Talk, Uutiset, Bettina S sekä Spotlight. Testattavana oli siis sekä suomenkielistä että ruotsinkielistä aineistoa.

4 METODIT

Tutkielman kysymyksenasettelun käytännönläheisenä lähtökohtana oli luvussa 3.3. kuvattu Ylen hanke videoaineiston automaattisen sisällönkuvailun kehittämiseksi. Tutkimuskysymysten jäsentämiseen haettiin ideoita myös aiemmista tutkimuksista, joita esiteltiin edellä luvuissa 2 ja 3. Tältä osin olivat erityisen hyödyllisiä Dalakleidin ja Kumppaneiden (2011) tutkimus sekä Viljasen ja Mattilan (2017) artikkeli.

4.1 Tutkimuskysymykset

Tämän tutkimuksen tarkoituksena on selvittää, miten Ylen videoaineistojen sisällönkuvailun parissa työskentelevät henkilöt arvioivat sisällönkuvailun nykyistä toimivuutta ja millaisiin tuloksiin heidän mukaansa päästiin automaattisen sisällönkuvailun hankkeessa. Tutkimuksessa haettiin vastauksia seuraaviin kysymyksiin:

- 1) Miten nykyinen sisällönkuvailu toimii Ylessä?
- 2) Miten automaattinen sisällönkuvailun ohjelma toimii Ylen videoaineistojen sisällönkuvailussa?
- 3) Mitä automaattisen sisällönkuvailun ohjelman tuottamasta sisällönkuvailusta puuttuu?
- 4) Miten automaattista sisällönkuvailua voitaisiin hyödyntää helpottamaan sisällönkuvailijoiden työtä?

Tutkielmaa ei ole tehty automaattisen sisällönkuvailun teknisen toteutuksen tutkimuksena, vaan tutkimuksen kohteena ovat käyttäjät ja heidän kokemuksensa automaattisen sisällönkuvailun parissa. Kyseessä on käyttäjäkeskeinen tutkimus, jonka kohteena on sisällönkuvailun työ.

Alkuperäisenä tutkimusaiheena oli tutkia sitä, miten automaattisen sisällönkuvailun ohjelman käyttöönotto onnistuu Ylellä ja miten sisällönkuvailijat suhtautuvat käyttöönot-

toon. Automaattista sisällönkuvailun ohjelmaa ei kuitenkaan otettu käyttöön sen ilmeistä puutteista johtuen. Näin ollen tutkimuksen näkökulmaa ja tutkimuskysymyksiä vaihdettiin kesken tutkimuksen.

Tämä ei kuitenkaan muuttanut alkuperäistä ajatusta siitä, että tutkimuksessa haluttiin selvittää nimenomaan sitä, miten käyttäjät suhtautuvat automaattiseen sisällönkuvailuun. Tästä johtuen tutkimuksen näkökulmaa oli suhteellisen helppo kohdentaa siihen, miten automaattinen sisällönkuvailu toimii nykyisin ja mitä puutteita siinä mahdollisesti on.

4.2 Haastatteluaineiston keruu

Tutkielman empiirinen aineisto koottiin haastattelemalla kolmea Ylen työntekijää. Heitä haastateltiin kerran noin tunnin ajan. Haastateltavat valikoitiin tutkimukseen työroolin perusteella. Sisällönkuvailun kaksi päätyöroolia, eli tuotantokoordinaattori sekä informatikko olivat edustettuina. Näiden henkilöiden lisäksi haastateltiin Ylen kehityspäällikköä. Alun perin tarkoituksena oli haastatella myös sisällönkuvailunhankkeen toteuttaneen yrityksen Valossan työntekijöitä, mutta valitettavasti haastattelua ei saatu järjestymään.

Haastateltavat olivat kaikki pitkään Ylellä ja varsinkin sisällönkuvailun parissa työskennelleitä henkilöitä, jotka olivat mukana testaamassa automaattisen sisällönkuvailun ohjelmaa. Näin ollen kaikilla haastateltavilla oli hyvä käsitys sisällönkuvailun parissa työskentelystä Ylellä ja mitä vaikutuksia automaattisella sisällönkuvailun ohjelmalla olisi ollut, jos se olisi otettu Ylellä käyttöön.

Kun halutaan, että tutkittava asia konkretisoituu inhimillisen kokemuksen kautta, ovat laadullisen tutkimuksen menetelmät sopivia tiedon keräämiseen. Seuraamalla haastattelun periaatteita, joissa annetaan tutkimuksen kohteen määritellä tutkimusmenetelmä, laadullisesta tutkimuksesta tulee objektiivinen tapa tutkia inhimillistä maailmaa. (Brinkmann & Kvale, 2005). Tämän vuoksi aineistonkeruun menetelmäksi valikoitui haastattelu. Aineistoa haluttiin kerätä tutkimuskohteen mukaan. Koska tutkimuksen kohteena oli nimenomaan henkilöt ja heidän inhimillinen kokemusmaailmansa, oli metodina haastattelu luonteva valinta tutkimusaineiston keräämiseen.

Haastattelut päätettiin suorittaa puolistrukturoituina, jota kutsutaan tässä tutkimuksessa Hirsjärveä ja Hurmetta mukaillen teemahaastatteluksi. Hirsjärven ja Hurmeen (2015, 47) mukaan teemahaastattelu on puolistrukturoitu menetelmä, joka pohjautuu Mertonin, Fiskin ja Kendallin kuvaukseen kohdennetusta haastattelusta. Kvalen (2006, 481) mukaan haastattelun tarkoitus on antaa ääni tavalliselle ihmiselle, jotta he voivat ilmaista itseään vapaasti omilla sanoillaan ja avautumaan tiiviissä yhteistyössä haastattelijan kanssa. Haastattelussa pyritään ymmärtämään maailma kohteen näkökulmasta. Tähän myös tässä haastattelussa pyrittiin. Tarkoitus oli antaa ääni niille, jotka olivat työssään kohdanneet uuden tekemisen tavan. Haastattelussa heillä oli mahdollisuus omin sanoin kertoa, miltä automaattisen sisällönkuvailun testaus tuntui. Tuomen ja Sarajärven (2018) mukaan teemahaastattelu viitekehyksenä toimii tieto, joka tutkittavasta aiheesta on etukäteen (Tuomi, Sarajärvi 2018). Koska tutkimuksen tekijällä oli etukäteen kerättyä tietoutta aiheesta, hänen oli mahdollista suorittaa teemahaastattelu siihen nojaten.

Haastattelu myös suoritettiin niin, että paikalla oli vain haastattelijä sekä haastateltava joten heidän välilleen pystyttiin muodostamaan luottamuksellinen tunnelma. Tutkimuksessa käytyä haastattelua ei voida kuitenkaan kutsua haastattelijan ja haastateltavan väliseksi keskusteluksi. Tämä johtui siitä, että haastattelu kontrastina keskustelulle yhtäältä etsii ymmärrystä, jonka haastateltava voi tarjota ja toisaalta palvelee haastattelijan tiedontarvetta (Kvale 2006, 483).

Koska haastattelijalla oli selkeä tiedontarve, oli haastattelu syytä kohdentaa siihen. Kohdennetun haastattelun tunnistaa neljästä asiasta. Ensinnäkin haastateltavat ovat kaikki kokeneet samankaltaisen tai yhtenevän tilanteen. Toiseksi tutkimuksen laatija on tutkinut alustavasti tutkittavan aiheen perusteet. Kolmanneksi tutkimuksen tekijä on laatinut tutkimukselleen haastattelurungon (ks. tämän tutkielman liite 1). Neljänneksi tutkimus kohdentuu haastateltavien omiin kokemuksiin tutkittavasta aiheesta. (Hirsjärvi & Hurme 2015, 47). Teemahaastattelu valikoitui metodiksi, ennen kuin oli tietoa siitä, että automaattisen sisällönkuvailun ohjelma oli liian keskeneräinen tuotantokäyttöön otettavaksi ja tutkimuksen painopistettä tulisi muuttaa. Koska haastattelut oli jo sovittu ja haastattelurunko laadittu, haastattelut suoritettiin alkuperäisen suunnitelman mukaisesti. Haastattelut tehtiin loppuvuodesta 2016 sekä alkuvuodesta 2017. Koska haastatteluiden aikana tiedettiin jo, ettei automaattisen sisällönkuvailun ohjelmaa tulla ottamaan käyttöön, haastattelut suoritettiin hieman puolistrukturoitua haastattelua vapaammin. Kaikissa haastatteluissa käytettiin samaa haastattelurunkoa, mutta haastatteluissa keskustelun annettiin polveilla

yli kysymysten. Näin saatiin laajempi näkemys käyttäjien tunteista ja tarpeista sekä nykyiseen sisällönkuvailuun että automaattiseen sisällönkuvailuun kohdentuen. Haastattelurunko toimi näin ollen pikemminkin haastattelun taustatukena.

Teemahaastattelu metodina sopii vapaamuotoisempaan haastatteluun hyvin, sillä se on lähempänä strukturoimatonta kuin strukturoitua haastattelua. Teemahaastattelua kuitenkin kutsutaan puolistrukturoiduksi, koska haastattelun teema on kaikille sama. (Hirsjärvi, Hurme 2015, 48).

Haastattelussa pyrittiin myös huomiomaan etiikka mikrotasolla. Brinkmann ja Kvale (2005) jakavat haastatteluiden etiikan mikro- ja makrotasolle. Mikrotasolla huomioidaan se, että haastateltavia kohdellaan asianmukaisella tavalla. Taataan haastattelun luottamuksellisuus sekä informoidaan haastateltavia tutkimuksen luonteesta. Lisäksi kerrotaan, että haastateltava voi vetäytyä haastattelusta niin halutessaan. Haastattelijan on myös huolehdittava siitä, että haastattelijalle ei synny ikäviä jälkiseurauksia haastattelusta.

Tutkimuseettisen neuvottelulautakunnan laatimien eettisten periaatteiden kolme osa-aluetta otettiin haastatteluissa huomioon.

- 1) Tutkittavan itsemääräämisoikeuden kunnioittaminen
- 2) Vahingoittamisen välttäminen ja
- 3) Yksityisyys ja tietosuoja

(Tutkimuseettinen neuvottelukunta 2009).

Tutkimuseettinen neuvottelukunta (2009) mukaan tutkittavan itsemääräämisoikeuden huomioimisessa tärkeää on, että haastateltava ymmärtää haastattelun olevan vapaaehtoista. Haastateltavalla tulee myös olla ennen haastattelun alkua riittävät tiedot haastattelusta. Edellä mainitut seikat liittyvät haastateltavan kunnioittavaan kohteluun. Kunnioittavan kohtelun kautta halutaan välttää minkäänlaisen vahingon tuottaminen haastateltaville. Yksityisyyden ja tietosuojan vaatimus tässä tutkimuksessa on toteutettu siten, että haastateltavien henkilöiden nimet eivät ole tutkimuksessa näkyvillä, vaan heihin viitataan ainoastaan heidän työnimikkeidensä kautta.

4.3 Affinity diagram -metodi

Tutkimuksessa koettiin tärkeäksi se, ettei tutkimus kohdentuisi vain yhteen käyttäjään tai yhteen käyttäjärooliin. Sovellus, kuten tässä tutkimuksessa tutkittu automaattisen sisällönkuvailun ohjelma, suunnitellaan monelle käyttäjäryhmälle. Tämän vuoksi tutkimuksessa on syytä ottaa huomioon tutkimuksessa esiintyvien yksilöiden lisäksi myös työpaikan rakenne sekä käyttäjien varianssi. Affinity diagram on laajalti tunnettu käyttäjälähtöisen suunnittelun metodi suurten tietomassojen käsittelyyn. Affinity diagrammin avulla voidaan selvittää esimerkiksi työpaikan avaintekijät, käyttäjien päämäärät ja kuinka ne saavutetaan, kipukohdat sekä käytetyt työkalut. (Beyer 2010, 30).

Affinity diagram –metodin nimestä on löydettävissä suomennoksia. Käytössä on esimerkiksi Karjalaisen (2007) käyttämä ryhmittelykaavio sekä Järvisen & Waltherin (2011) tutkielmassaan käyttämä samankaltaisuuskaavio. Tässä tutkielmassa käytetään kuitenkin englanninkielistä versiota affinity diagram, koska se on metodin yleisimmin käytetty nimi.

Affinity diagram ei ole informaatiotutkimuksessa laajalti käytössä oleva metodi. Sen juuret ovat Jiro Kawakitan luomassa, johdon avuksi tehdyssä päätöksentekoon liittyvässä KJ-metodissa (Scupin 1997, 233). KJ-metodi kehittyi Kawakitan analysoidessa etnografisia aineistoja. Scupin artikkelissa on suora lainaus siitä, mitä Kawakita ajatteli, kun hänen edessään oli suuri massa aineistoa, joka piti käsitellä. Kawakitan kerrotaan ymmärtäneen valtavaa aineistomassaa katsellessaan, että aineistojen erilaiset järjestykset antavat mahdollisuuden nähdä ne eri lailla. Aineiston järjestäminen vaikuttaa siis siihen mitä ne kertovat. (Scupin 1997).

Tässä tutkimuksessa oli ehditty jo sopia loppukäyttäjien kanssa aineistonkeruu menetelmäksi teemahaastattelu sekä muodostaa haastatteluiden runko kun selvisi, että tutkimuksen suuntaa tulisi vaihtaa pois käyttöönottotutkimuksesta kohti käyttäjäkokemustutkimusta. Tämä tarkoitti käytännössä sitä, että aineisto piti järjestää ja analysoida tavalla, joka kertoisi käyttäjien näkökulmasta sisällönkuvailusta ja varsinkin automaattisesta sisällönkuvailusta jotakin uutta.

Affinity diagram –metodi oli tutkimuksen tekijälle entuudestaan tuttu käytettävyystudiumuksen puolelta, jossa sitä käytetään analysoimaan esimerkiksi loppukäyttäjähaastatteluita. Tarkoituksena oli selvittää, mitkä olivat käyttäjien todelliset tarpeet automaattisen

sisällönkuvailun ohjelman osalta ja miten he kokivat sisällönkuvailun automaattisen sisällönkuvailun testauksen aikana. Näiden asioiden selvittämiseen affinity diagram on metodina hyvä, koska menetelmää käytetään, kun halutaan selvittää henkilöiden todelliset tarpeet. Käyttäjiltä ei voi Holtzblattin ja Beyerin (2015, 11) mukaan vain kysyä mitä he tarvitsevat, koska se olisi kuin kivien poimimista rannalta. Viittaus kivien poimimiseen voidaan liittää siihen, että kysyttäessä tarpeet voivat olla yksittäisiä toivomuksia tai huomioita. Kiviä rannalta poimimalla kerätään vain hajanainen joukko joitakin huomioita. Nämä yksittäiset huomiot eivät korreloi keskenään, eivätkä anna kokonaiskuvaa yksilön tarpeista, toiveista tai ongelmista.

Ongelma on hajanaisten tarpeiden ilmaisemisen lisäksi siinä, etteivät yksilöt välttämättä tiedä mihin kaikkeen teknologia pysyy, eivätkä näin ollen pysty määrittelemään, mitä siltä toivovat. Suurimpana ongelmana voidaan kuitenkin pitää sitä, etteivät yksilöt ole selvillä siitä, mitä he oikeastaan tekevät. Kun samoja toimintoja toistetaan joka päivä, niistä tulee tiedostamattomia, eikä rutiineja ole helppo artikuloida. Yksilöt yleensä kyllä pystyvät kertomaan yleisellä tasolla siitä, mitä heidän jokapäiväiseen toimintaansa kuuluu tai kertomaan kriittisistä ongelmista, joita he kohtaavat järjestelmän parissa. Tavallisesti yksilöt osaavat esimerkiksi kertoa, mikä saa heidät turhautumaan heidän käyttäessään jotakin järjestelmää. Vaikeampaa kuitenkin on kertoa sisäisistä motivaatioista, kuten ilmaista identiteettiään. Tämän menetelmän tarkoituksena on nostaa esille myös niitä puolia, joita henkilö ei osaa artikuloida selkeästi. (Holtzblatt & Beyer 2015, 11).

Elämä on monimutkaista ja yksityiskohtaista, joten kenttätutkimuksen tulosten jakaminen voi olla hankalaa. Kuitenkin jopa pieni projekti hyötyy rikkaasta käyttäjädatasta. Käyttäjätiedon avulla ymmärrämme paremmin käyttäjien maailmaa. Metodin tavoitteena on auttaa järjestelmien kehittäjiä sisäistämään tuotteen käyttäjien maailma eli muodostaa siltä käyttäjien ja järjestelmien kehittäjien välille. (Holtzblatt & Beyer 2015, 21-22).

Aineiston konsolidoiminen affinity diagram –metodin avulla saattaa yhteen kaiken tutkittavina olleista henkilöistä kerätyn datan yhteen yhdenmukaiseen näkymään, josta on havaittavissa yleisimmät käytösmallit ilman, että hävitetään yksilöistä johtuvaa varianssia. (Holtzblatt & Beyer 2015, 23). Merkintöjen konsolidoiminen affinity diagram -metodin avulla mahdollistaa koko tutkittavien henkilöiden ryhmän nostamat ongelmat kohdat, eikä tutkimus spesifioitu vain yhteen yksilöön (Holtzblatt, Wendell & Wood 2005,

159). Kvantitatiiviset menetelmät nostavat aineistoista esiin yleensä vain muutamia löydöksiä. Muutamien parhaiden löydösten esille tuominen datasta kuitenkin hävittää aineiston alkuperäisen rikkauden. Tämän metodin tavoitteena on ottaa erilainen lähestymistapa aineistojen jäsentelyyn. Rikas ja yksityiskohtainen aineisto järjestellään niin, että aineisto paljastaa avainongelmat, yleisimmät rakenteet sekä kaavat, joita noudatetaan. (Holtzblatt & Beyer 2015, 23). Affinity diagram -metodi nostaa ylös ongelmia ja oivalluksia seinän kokoiseen hierarkkiseen diagrammiin (Holtzblatt, Wendell & Wood, 2005, 159).

Kerätty data muunnetaan yksittäisiksi merkinnöiksi. Jokainen yksittäinen merkintä kirjoitetaan omalle lapulleen, lappu sisältää aina vain yhden idean (Holtzblatt, Wendell & Wood 2005, 167). Jokainen lappu on luettavissa yksinään, eikä lapun ymmärtämiseen tarvita muuta kontekstia. Nämä omille lapuilleen jaetut merkinnät edustavat koko kerättyä dataa (Holtzblatt, Wendell & Wood 2005, 159). Laput sekoitetaan satunnaiseen järjestykseen (Holtzblatt & Beyer 2015, 24). Affinity diagrammia rakennetaan alhaalta ylöspäin, ryhmittelemällä yksittäisiä merkintöjä, jotka paljastavat avainteemoja kerätystä datasta. Datan annetaan ehdottaa luokitteluita mieluummin kuin, että aloitettaisiin ennalta määrättyistä luokitteluista. (Holtzblatt, Wendell & Wood, 2005, 159).

Monissa aineistojen käsittelyyn tarkoitetuissa koodausmetodeissa kysymykseksi nousee se, miten aineistojen luokat määräytyvät. Yleensä laadullisen tutkimuksen koodausta tehtäessä hyödynnetään induktiivista lähestymistapaa, jossa tutkija ikään kuin sukeltaa datan sisälle ja muodostaa kategoriat datasta. Toisaalta deduktiivinen menettely ei ole harvinaisen. Tässä menetelmässä tutkijat ottavat esimerkiksi kirjallisuudesta tai aiemmasta tutkimuksesta ennalta määriteltyjä kategorioita. Aineiston käsittelyn edetessä luokkia karsitaan tai lisätään tarpeen mukaan. Tämän tyyppinen laadullinen tutkimus toimii parhaiten tarkasteltaessa oletuksia tai valmiita konsepteja. Käyttäjän tarpeisiin tai käyttökokemukseen liittyvä tutkimus on yleensä induktiivista. Prototyyppihin liittyvä tutkimus sen sijaan on usein deduktiivista. (Stickdorn et al. 2018, 113-114). Vaikka tämä tutkimus koskeekin prototyyppin testausta, tavoitteena oli selvittää käyttäjien tarpeita sekä käyttökokemusta. Siksi aineiston käsittelyyn valikoitui Stickdornin et al. (2018) jaon mukaan induktiivinen eikä deduktiivinen menetelmä.

Jossain määrin tutkielmassa käytettyä tutkimusmetodia voidaan pitää aineistolähtöisenä analyysinä. Tuomen ja Sarajärven mukaan aineistolähtöisissä analyysissä kohteet vali-

taan tutkimuksen tarkoituksen ja tehtävänasettelun mukaan. Analysoitavia kohteita aineistosta ei ole siis valittu ennen analyysin suorittamista. Idea on siinä, että aikaisempi tietämys aiheesta ei ohjaa analyysiä, eikä vaikuta siihen, mihin loppupäätelmään analyysissä tullaan. Aineistolähtöinen analyysi ei ole ongelmaton menetelmä, jonka ongelmat on syytä tunnistaa. Aineistolähtöisessä analyysissä tutkijan pitäisi yrittää unohtaa tietonsa tutkittavasta aiheesta. Voidaan asettaa kyseenalaiseksi, voiko tutkimuksen tekijä todella unohtaa aikaisemmat tietonsa aiheesta, ja pystyä näin hahmottamaan aineistosta vain ja ainoastaan aineistosta kumpuavia havaintoja. Tutkimuksessa on otettava huomioon se, että aikaisempi tieto ja tunnettu teoria todennäköisesti jossakin määrin vaikuttavat tutkimuksen tuloksiin. Tärkeää on kuitenkin huomata, että tutkimuksen tarkoituksena ei ole testata aikaisempia teorioita vaan löytää aineistosta uusia havaintoja tutkittavasta aiheesta. (Tuomi & Sarajarvi, 2018).

4.4 Empiirisen aineiston analyysi

Kaikki kolme haastattelua nauhoitettiin ja nauhoitteet litteroitiin mahdollisimman tarkasti. Litteroidut haastattelut jaettiin affinity-merkinnöksi affinity diagram –metodin mukaisesti. Käytännössä haastattelut jaettiin lappuiksi niin, että jokainen affinity-merkintä oli luettavissa itsenäisesti ilman kontekstia ja sisältäen vain yhden asian kuvaten kuvasta 1 on havaittavissa.

E-07	Mähän en kirjoita sitä mitä ihmiset sanoo vaan se on niinkun kuvan määrittelemisen, että mitä siinä kuvassa on
H-52	toki olispa hyvä, että meidän järjestelmän integraatio toimis paremmin

Kuva 1. Kaksi esimerkkiä affinity-merkinnöistä.

Jokaiseen lappuun lisättiin identifikaation mahdollistava tunnus, joka sisälsi haastateltavan nimikirjaimen sekä liukuvassa järjestyksessä annetun järjestysnumeron (kuva 1). Järjestysnumero annettiin aina haastateltavakohtaisesti, joten jokaisen ensimmäinen lappu sai numerokseen numeron yksi. Kaikkiaan kolmesta haastattelusta muodostettiin 447

kappaletta affinity-merkintöjä. Kaikki affinity-merkinnät leikattiin erillisiksi lapuiksi ja laput sekoitettiin keskenään.

Lappujen sekoittamisen jälkeen jokainen niistä poimittiin satunnaisessa järjestyksessä ja liimattiin seinälle. Saman sisältöiset laput liimattiin lähelle toisiaan, kuitenkin pyrkien välttämään sitä, ettei samalta henkilöltä järjestysnumeroitaan lähellä olevia lappuja laiteta samaan rivistöön. Aivan täysin tältä ei voitu välttyä, mutta pyrkimys oli erottaa liian lähekkäin olevat affinity-merkinnät toisistaan. Kun laput alkoivat muodostaa selkeitä kokonaisuuksia, ne otsikoitiin. Kuvassa 2 näkyy seinälle liimattuja affinity-merkintöjä sekä niiden muodostamia rivejä, joille on annettu ensimmäisen tason otsikoita.



Kuva 2. Affinity diagram alkaa muodostua seinälle.

Yhden ensimmäisen tason otsikon alle pyrittiin laittamaan noin 4-8 lappua (kuva 2). Kuitenkaan kaikkien otsikoiden kohdalla tähän ei päästy. Lappujen kuvitteellinen maksimiluku auttoi hahmottamaan, milloin oli syytä muodostaa yhden ryhmän sijasta kaksi lappuryhmää. Kaikkiaan ensimmäisen tason otsikoita muodostettiin 80 kappaletta. Kun ensimmäisen tason otsikoita alkoi muodostua seinälle, niissä oli havaittavissa samankaltaisuuksia, nämä sijoitettiin seinällä lähekkäin ja näin muodostuneelle uudelle ryhmälle annettiin toisen tason otsikko. Toisen tason otsikoita aineistosta muodostui kaikkiaan 31 kappaletta. Tavoitteena oli saada toisen tason otsikon alle noin 2-5 ensimmäisen tason

otsikkoa ja niiden muodostamaa affinity-merkkintä ryhmää. Tämä toteutui kaikkien muiden paitsi yhden osalta. Toisen tason otsikoita siirrellessä seinälle ne alkoivat muodostaa seinälle myös omaa ryhmäänsä. Myös nämä toisen tason otsikoiden sekä niiden alla olevien otsikoiden alla olevat ryhmät saivat kolmannen tason otsikon. Kaikkiaan kolmannen tason otsikoita tästä aineistosta muodostui 11 kappaletta. Tavoitteena oli saada jokaisen kolmannen tason alle 1-5 toisen tason otsikkoa. Tämä saavutettiin. Kuvassa 3 nähdään kuinka kolmannen tason alapuolella on sekä toisen että ensimmäisen tason otsikoita ryhmiteltyinä.



Kuva 3. Kolmannen tason otsikko, sekä sen sisältävät alemman tason otsikot.

Huomioitavaa on, että lappujen asetteleminen ja otsikointi eivät tapahtuneet järjestyksessä. Affinity -merkinnoistä muodostuneita ryhmiä otsikoitiin aina kun ryhmä täyttyi tai

kun metodin käyttäjä ymmärsi niiden muodostavan ajatuksen. Kaikkia affinity -merkin-
töjä ei ollut vielä laitettu seinälle, kun ensimmäisiä kolmannen tason otsikoita oli jo muo-
dostunut.

Affinity diagrammin valmistuttua siitä muodostettiin taulukko (liite 2). Affinity diagram-
-taulukon valmistumisen jälkeen itse analysointivaihe alkoi. Jokaisen kolmannen otsikon
muodostama taulukon osa analysoitiin erikseen. Käytännössä analysoija tutki, mitä affi-
nity-merkin-
töjä otsikoiden alle sijoittui ja perusteli niiden avulla affinity diagrammin
muodostamisen vaiheessa syntyneet otsikot. Näin aineistosta syntyneet havainnot pystyt-
tiin kattavasti alkuperäisen haastatteluaineiston avulla analysoimaan. Koska aineisto oli
lajiteltu otsikoiden alle välittämättä kronologisesta järjestyksestä tai siitä kuka haastateltu
huomion oli sanonut, oli useimmiten yhden otsikon alla kaikkien haastateltujen kom-
mentteja. Usein affinity-merkin-
töjä oli myös useita yhdeltä haastatellulta, mutta kom-
mentit oli sanottu haastattelun eri vaiheissa. Affinity diagram –metodin avulla aineistoa
voitiin analysoida vertaillen sitä, mitä kukin haastateltava oli aiheesta sanonut. Analyys-
sissa pohdittiin, olivatko haastatellut aiheesta saman mielisii, vai oliko heidän välillä
eroavaisuuksia. Analysoijan ei tarvinnut analyysivaiheessa huolehtia siitä, tuleeko ai-
neisto vastaamaan tutkimuskysymyksiin, koska aineisto oli kerätty teemahaastattelu- me-
todin avulla.

Affinity diagram oli metodina ja haastatteluiden analyysin apuvälineenä tässä tutkiel-
massa kohtalaisen raskas menetelmä. Tämä johtuu siitä, että koko aineistoa pidetään ana-
lyysin aikana aktiivisesti mukana. Koska affinity diagram –metodi on kehitetty suurien
aineistomassojen käsittelyyn, sen kantavana ajatuksena on se, ettei yksittäisiin affinity-
merkin-
töihin tarvitse palata enää sen jälkeen, kun otsikkoihin kirjatut päähavainnot on
löydetty. Tämä tutkielma on kuitenkin luonteeltaan tarkempi, joten analyysivaiheessa
myös affinity-merkinnät olivat aktiivisen tarkastelun kohteena.

Kun koko affinity diagram –taulukko oli analysoitu, siitä muodostettiin tutkielman luvut
niin, että jokainen kolmannen tason otsikko esiintyy tutkielman luvun otsikkona. Luvun
sisään on kirjoitettu analyysi. Koska tutkielma pohjaa vahvasti affinity diagrammista
muodostuneeseen taulukkoon, on lukua koskeva taulukon osa sijoitettuna kyseisen luvun
loppuun. Näin lukijan on helppo nähdä, millaisesta taulukonosasta analyysi on muodos-
tunut.

5 SISÄLLÖNKUVAILUN KÄYTÄNNÖT YLELLÄ

Tässä luvussa siirrytään empiiristen tulosten tarkasteluun. Erityisesti keskitytään vaatimuksiin, joita multimedialliset aineistot asettavat sisällönkuvailulle.

5.1 Sisältöjä pitäisi voida kuvailla ilman formaalia koulutusta

Sisällönkuvailuja tekevät Ylellä niin tuotantokoordinaattorit kuin informaattikotkin. Informaattikoilla on työtä vastaava korkeakoulututkinto tai pitkän työuran mukanaan tuoma kokemus. Tuotantokoordinaattoreilla ei ole sisällönkuvailun koulutusta, vaan usein heidän koulutuksensa on medianomin amk-tutkinto tai muu media-alantutkinto. (Yle, 2016). Ei voida siis olettaa, että sisällönkuvailija on alan ammattilainen (taulukko 1).

Sisällönkuvailun prosessi etenee niin, että tuotantokoordinaattorit tekevät sisällönkuvailut ja informaattikot korjaavat tehdyt kuvailut (taulukko 1). Kuvailuja tarkistamalla informaattikot huolehtivat, että tarvittava sisällönkuvailu on saapunut tuotantokoordinaattoreilta toimituksesta arkistolle. Tarvittaessa informaattikot korjaavat sekä täydentävät kuvailuja. He myös antavat palautetta ja opastavat tuotantokoordinaattoreita sisällönkuvailuun liittyvissä asioissa. Monivaiheisen sisällönkuvailun prosessin etuna on se, että ensimmäisen sisällönkuvailun tekee ohjelmaprosessissa mukana ollut henkilö, eli tuotantokoordinaattori. Tällöin sisällönkuvailun ensimmäisen vaiheen tekee henkilö, jolla on tietoa esimerkiksi kuvauspaikoista sekä ohjelman tekemiseen osallistuneiden henkilöiden nimistä. Informaattikko kertoo haastattelussa, että hän harvoin aloittaa sisällönkuvailua tyhjistä. Tämä voidaan tulkita niin, ettei informaattikon työnkuvaan Ylellä ainakaan valitsevasti kuulu ohjelmien kuvaileminen alusta alkaen. Haastattelussa selvisi myös, että arkistossa kuvaillaan ainoastaan vanhoja aineistoja alusta.

Tuotantokoordinaattorit tekevät suurimmalle osalle aineistoja ensimmäiset sisällönkuvailut. Sisällöt kuvaillaan heidän toimestaan seuraavasti. Ohjelma jaetaan manuaalisesti segmentteihin eli osuuksiin. Esimerkiksi uutisissa yksi segmentti sisältää aina yhden uutisjutun. Segmenttien avulla sisällönkuvailu saadaan kohdennettua oikeaan kohtaan ohjelmassa. Joissakin ohjelmissa, esimerkiksi dokumenteissa, jotka kertovat selkeästi vain yhdestä aiheesta, yksi segmentti saattaa sisällönkuvailussa kattaa koko ohjelman. Kuvasta

4 on nähtävissä, kuinka Metro-järjestelmässä kirjoitetaan aihe, kuvankuvailu, tekijät, kuvälähteet ja haastateltavat sisällönkuvailuun.



Kuva 4. Näkymä Metro-järjestelmän sisällönkuvailusta.

Sisällönkuvailuun kirjoitetaan ne tekijät, joille on ohjelmasta muodostunut tekijänoikeus, sekä luetellaan myös ne henkilöt, jotka ovat oleellisesti vaikuttaneet ohjelma syntymiseen. Jokainen tekijä kirjoitetaan omalle rivilleen etunimi sukunimi –muotoon (kuva 4).

Ohjelman aiheeksi kirjoitetaan lyhyt kuvaus ohjelmasta. Jos kyseessä olisi esimerkiksi Suomen ja Ruotsin välinen jääkiekko-ottelu jääkiekon maailmanmestaruuskilpailuissa kirjattaisiin aihe seuraavasti: "Jääkiekon MM-kilpailut Slovakiassa loppuottelu Suomi – Ruotsi (6-1)". Näin ollen kuvan erittelyssä kerrotaan lyhyesti, kuka kilpailee ja missä lajissa. Jos kyseessä on makasiiniohjelma aiheeksi riittää esimerkiksi: "A-studio: ajankohtainen keskusteluohjelma". Yksittäisten juttujen aiheet kuvataan tarkemmin kuvasisällön erittelyn yhteydessä. Juontajat ja studiovieraat kirjoitetaan titteli Etunimi Sukunimi muodossa, esimerkiksi: "pääministeri Juha Sipilä". Keskusteluohjelmista kirjoitetaan tiivis erittely, josta on luettavissa haastateltavien ja keskustelijoiden nimet. Näin ollen aiheena voisi olla esimerkiksi: "Pekka Haaviston haastattelu", tekijänä toimittajan nimi "Niko Hannunen (toimittaja)", kuvasta kirjoitetaan mitä siinä tapahtuu eli "Haastattelussa presidenttiehdokas Pekka Haavisto". Itse haastattelun sisältö tiivistetään kuvailussa muutama sanaan. Silloin jos sisältö koetaan erityisen tärkeäksi, se voidaan siteerata sanasta sanaan. Sisällönkuvailu tehdään käyttäen hyvää suomen kieltä ja kokonaisia lauseita. (Halonen, Kajo, 2015).

Sisällönkuvailu tehdään Ylessä noudattaen melko tarkkoja ohjeita, jotta sisällönkuvailut pysyisivät mahdollisimman tasalaatuisina. Kohtalaisen tarkkoista ohjeista huolimatta, ihmillisistä syistä johtuen, sisällönkuvailujen laadussa on vaihteluita. Gotohin ja Khanin (2017, 244-246) mukaan ihmiset ovat hyvin subjektiivisia kuvaillessaan videoita. Heidän mukaansa siis kaksi ihmiskuvailijaa tuottavat kaksi hyvin erilaista kuvausta samasta videoaineistosta. Siinä missä toinen sisällönkuvailija voi nähdä harvinaista kuvaa esimerkiksi uhanalaisesta eläimestä, toinen sisällönkuvailija ei välttämättä tiedä eläimen olevan uhanalainen ja jättää tämän seikan kokonaan kuvailusta pois. Tätä ongelmaa voidaan yrittää selittää erilaisilla kuvailuohjeistuksilla, kuvailusäännöillä tai vaikkapa avainsanoja käyttämällä.

Lisäksi vaihtuvuutta sisällönkuvailujen tasalaatuisuuteen Ylellä tuo se, että tuotantokoordinaattoreilla, jotka sisällönkuvailun tekevät, ei ole formaalia koulutusta asiaan (taulukko 1). Koulutuksen puute voi näkyä esimerkiksi siinä, että sisällönkuvailujen tärkeyttä ei aivan ymmärretä, tai sisällönkuvailua ei osata tehdä niin, että se olisi tiedonhaun kannalta mahdollisimman optimaalista.

Haastatteluissa selvisi myös, että Ylellä tapahtuva sisällönkuvailu kohtaa kieliongelmia. Ylen tuotantokoordinaattoreista kaikki eivät puhut äidinkielenään suomea, mutta kaikki sisällönkuvailut kuitenkin tuotetaan suomen kielellä. Haastattelussa informaatikko kertookin, että sisällönkuvailussa ilmenee jonkin verran kieliongelmia, joten joitakin sisällönkuvailuja on korjattava enemmän kuin toisia.

Taulukko 1. Sisältöjä pitäisi voida kuvailla ilman formaalia koulusta asiaan

Ei voida olettaa, että sisällönkuvailija on alan ammattilainen	Tuotantokoordinaattorit kuvailevat, informaatikot korjaavat
	Sisällönkuvailua pitäisi voida tehdä ilman asiantuntijuutta
	Sisällönkuvailun parissa työskentelee monenlaisia ammattilaisia

5.2 Sisällönkuvailijana haluan tehdä työni tehokkaasti

Sisällönkuvailun työ halutaan tehdä mahdollisimman tehokkaasti (taulukko 2). Samaa aineistoa ei haluta kuvailla pitkään tai montaa kertaa, koska työn kustannukset nousisivat. Sisällönkuvailu halutaan saada kuntoon nopeasti ja niin, ettei sisällönkuvailuun tarvitse enää palata.

Sisällönkuvailun työtä halutaan siis tehostaa ja yksi syy sisällönkuvailun tehostamiseen on kustannukset. Kuvailtavan videoaineiston massa on valtava ja jos kaikki aineistot haluttaisiin kuvailla ihmisen tekemänä työnä, vaatisi se paljon resursseja työn toteuttamiseksi. Koska sisällönkuvailijat ovat kallis resurssi, heidän aikaansa halutaan säästää (taulukko 2). Automaattisen sisällönkuvailun testauksen yksi tärkeimmistä tutkimuskohteista oli saada selville, voitaisiinko automaation avulla säästää sisällönkuvailijoiden aikaa tai helpottaa heidän työtään. Tähän ei kuitenkaan testissä saatu kattavaa vastausta. Automaattisen sisällönkuvailun ohjelma ei vielä toiminut tämän tutkimuksen aikana riittävällä tasolla, eikä voida poissulkevasti sanoa nopeuttaisiko automaattinen sisällönkuvailu sisällönkuvailijoiden työtä.

Tutkimuksessa selvisi kuitenkin, että automaattisella sisällönkuvailulla saattaisi olla joidakin etuja ihmisen tekemään sisällönkuvailuun verrattuna. Yksi näistä edusta olisi se, että automaattinen sisällönkuvailun ohjelma saattaisi, toisin kuin ihminen, kuvailla saman ohjelman useaan kertaan (taulukko 2). Jos oletetaan, että automaattinen sisällönkuvailun ohjelma olisi oppinut tunnistamaan enemmän objekteja kuvailujen välillä, olisi sisällönkuvailun uudelleen tekeminen ehkä mielekästä. Toinen vaihtoehtoinen käytätapa uudelleen kuvailuun voisi olla ajanmuuttuminen tehtyjen kuvailtujen välillä. Ajan muuttumisella tässä yhteydessä tarkoitetaan kuvailutietojen päivittymistä tai niiden täsmentymistä nykyajan vaatimuksia vastaavaksi. Kuten aiemmin mainittiin, ensimmäisen Yle Arkiston omakustanteiset videoaineistot on arkistoitu jo vuodesta 1957 alkaen, joten voidaan olettaa, että kuvailujen tarve ja käytätapa on vuosien varrella muuttunut. Enserin (1999) mukaan kuvailu, joka oli käypä kuvailun tekohetkellä, kuvastaa vain silloista näkemystä kuvailun tarpeista. Voidaan siis päätellä, että aineistoja olisi syytä kuvailla uudelleen, jotta kuvailuihin saadaan istutettua kulloisenkin aikakauden tarpeet kuvailulle.

Automaattista sisällönkuvailua voisi siis hyödyntää aineistojen uudelleen kuvailuun, toisin kuin ihmisiä, koska sisällönkuvailija ei lähtökohtaisesti halua tehdä samaa kuvailua

enää uudelleen (taulukko 2). Kehityspäällikkö pohtii haastattelussa, että tämä saattaa olla yleisempikin ilmiö ihmisillä:

”Onks se joku yleisempikin ilmiö, että sitten kun se joku asia lasketaan käsistä, niin se on ikään kuin kohtalaisen valmiiksi kuitattu”

Ihmislähtöinen sisällönkuvailu tähtää kehityspäällikön mukaan siihen, että samaan ohjelmaan ei tarvitse enää palata. Kun ohjelma on kuvailtu kertaalleen niin hyvin, että sen voidaan olettaa olevan valmis ja ihmisen luettavissa, on kuvailu lopullisesti riittävä ja valmis. Vaikka informaatikot tarkistavat osan ohjelmista, jonka tuotantokoordinaattorit ovat kuvailleet (taulukko 2), eivät informaatikot silti korjaa tai muuta jokaista tarkistamaansa kuvailua. Jos tuotantokoordinaattorit kuvailevat ohjelmat riittävän hyvin, eivät informaatikot kuvaile ohjelmaa enää uudelleen. Informaatikko kertoo, että jos hän toteaa kuvailun olevan riittävällä tasolla, käyttää hän ajallisesti yhdenohjelman kohdalla aikaa noin minuutin verran.

Ylellä työprosessi on tällä hetkellä kehityspäällikön mukaan suunniteltu niin, että sisällönkuvailijoiden ei tarvitse työssään moneen kertaan palata saman ohjelman sisällönkuvailuun. Tähän samaan prosessiin pyrittiin pääsemään myös automaattisen sisällönkuvailun testissä. Testissä onkin havaittavissa, että sisällönkuvailussa on yhteneviä prosessin vaiheita riippumatta siitä tekeekö kuvailun kone tai ihminen (taulukko 2). Yksi yhtenevä prosessin vaihe on esimerkiksi se, että sekä ihmisen että koneen tuottamaa sisällönkuvailua joudutaan korjaamaan. Informaatikko kertoi, että A-studio on videoaineistona sellainen ohjelma, jossa on usein aiheen puolesta korjattavaa. Tämä johtuu informaatikon mukaan A-studion aiheiden moninaisuudesta.

Informaatikko kertoo, että hänen työkuvaansa kuuluu olennaisena osana toimituksen puolella tehtyjen sisällönkuvailujen tarkastaminen, täydentäminen ja arviointi. Tämän työvaiheen ei tarvitsisi muuttua, teki ensimmäisen sisällönkuvailun ehdotuksen toimituksessa työskentelevä tuotantokoordinaattori tai automaattisen sisällönkuvailun ohjelma, sillä sisällönkuvailu käy läpi ihmisen tekemän tarkistusprosessin tälläkin hetkellä (taulukko 2). Eroa työprosessien välille tulee kehityspäällikön mukaan siinä, että ihmisen tekemää sisällönkuvailua ei tarvitse validoida.

Toinen yhteneväinen työprosessin vaihe on se, että sekä automaattinen sisällönkuvailun ohjelma että ihminen sisällönkuvailijana tarvitsevat ohjeet siihen, minkälaista sisällönkuvailua halutaan tuottaa. Sisällönkuvailijoille on laadittu Ylen sisäinen sisällönkuvailun

ohjeistus, jotta sisällönkuvailu pysyisi mahdollisimman tasalaatuisen hyvänä. Automaattiselle sisällönkuvailun ohjelmalle ohjeet tulee kehityspäällikön mukaan antaa parametreina. Parametrien avulla voidaan määritellä millaista metadataa ohjelma tuottaa. Kehityspäällikkö totesikin, että sekä ihmisen tuottama sisällönkuvailu että automaattisen sisällönkuvailun tuottama sisällönkuvailu, ovat muokattavissa olevia asioita riippuen siitä, miten niiden tekeminen ohjeistetaan (taulukko 2).

Jotta sisältöjen kuvaileminen olisi mahdollisimman tehokasta, sisällönkuvailijat haluaisivat hyödyntää kerran kirjattua tietoa mahdollisimman paljon. Työprosesseja voitaisiinkin tehostaa hyödyntämällä nykyistä paremmin jo kertaalleen syötettyjä tietoja. Järjestelmien välinen integraatio pitäisi siis toimia niin hyvin, ettei käyttäjän tarvitse huolehtia tietojen siirtymisestä järjestelmien välillä. Sisällönkuvailijat eivät ymmärrettävästi halua syöttää useita kertoja samoja tietoja eri järjestelmiin (taulukko 2). Järjestelmäintegraatio tehostaisi informaattikon mukaan työtä paljon. Esimerkkinä integraation mahdollistamasta tietojen käytöstä informaattikko antaa planssitekstit. Tekstit ovat jo olemassa olevaa tietoa toimitusjärjestelmässä siinä vaiheessa, kun automaattisen sisällönkuvailun ohjelma yrittää lukea planssitekstejä kuvapinnoista.

Taulukko 2. Aineistot eivät ole homogeenisiä ja tarvitsevat siksi erityyppisiä kuvailuja

Sisällönkuvailija haluaa tehdä työnsä tehokkaasti kerralla kuntoon	Sisällönkuvailijoilta halutaan säästää aikaa
	Sisällönkuvailija ei halua tehdä uudestaan samaa sisällönkuvailua
Sisällönkuvailussa on yhteneviä prosessinvaiheita riippumatta siitä tekeekö kuvailun kone vai ihminen	Niin ihmisenkin kuin koneenkin tuottama sisällönkuvailu joutuu korjaamaan
	Sekä kone, että ihminen tarvitsevat ohjeet siihen millaista kuvailua halutaan
	Sisällönkuvailu käy läpi ihmisen tekemän tarkistusprosessin tälläkin hetkellä
Haluan hyödyntää kerran kirjattua tietoa mahdollisimman paljon	En halua syöttää montaa kertaa samoja tietoja

	Jos tietoja on jo valmiina en halua syöttää niitä uudelleen
--	---

5.3 Aineistot eivät ole homogeenisiä

Sisällönkuvailussa on otettava huomioon rajatunkin aineistolajin sisäinen varianssi. Tuotantokoordinaattorin mielestä menetetään paljon, jos erityisen hyvälaatuisia ja harvinaisia maisemakuvia ei kuvailla riittäväällä tarkkuudella (taulukko 3). Maisemat voivat hänen mukaansa kadota kokonaan tai muuttua radikaalisti, joten videomateriaali olisi erityisen arvokasta. Erityislaatuisten aineistojen riittäväällä tarkkuudella tehty kuvailu on tuotantokoordinaattorin mielestä erityisen tärkeää myös siksi, että Yle on valtiollinen instituutio, jolla on velvollisuus säilyttää kansallisomaisuutta. Aineistojen säilyttämisestä ei kuitenkaan ole hyötyä, jos käyttäjät eivät löydä aineistoja. Maisemakuvat saattavat olla kulttuurihistoriallisesti merkittävää aineistoa. Enseri ja Sandomin (2002) mukaan aineiston kulttuurihistoriallista merkitystä ei voi yksiselitteisesti määritellä. Ensinnäkin arvokkaiden aineistojen määrittely on aina subjektiivista, koska kuvat lähettävät erilaisia viestejä eri ihmisille. Toiseksi on huomattava, ettei aina ole helppoa arvioida millainen kuva tai aineisto olisi mahdollisesti hyödyllinen tulevaisuudessa.

Aineiston kulttuurihistoriallisen merkityksen arvioinnin hankaluuden lisäksi näiden aineistojen kohdalla on toinenkin sisällönkuvailuun kohdentuva ongelma. Maisemakuvien kaltaiset kulttuurihistoriallisesti merkittävät aineistot eivät nimittäin tämän tutkimuksen mukaan ole se aineistomuoto, joka eniten määrittää Ylen sisällönkuvailun kehitystyötä (taulukko 3). Tuotantokoordinaattori kertoi, että sisällönkuvailun tarve on hyvin erilainen siellä, mistä kehitystyö tunnetaan Ylessä lähtevän, eli uutisaineistoista.

”Se on tässä ongelma mun mielestä kun tehdään sisällönkuvailuja niin uutiset ei sais aina olla se mikä määrittelee mitä me tehdään”

Informaatikko kertoo, että uutiset lähtökohtaisesti kuvaillaan tarkimmin ja huolellisimmin kuin mikään muu aineisto. Erityisen tarkasti kuvaillaan päivän pääuutislähetys, josta informaattikon mukaan huolehditaan siten, että siitä on kuvailtu kaikki.

Erilaiset aineistot siis saavat erilaista kohtelua sisällönkuvailussa (taulukko 3). Toiset aineistot kuvaillaan yleisemmällä tasolla ja toiset aineistot kuvaillaan erityistä tarkkuutta

noudattaen. Tuotantokoordinaattorin mielestä liian yleisellä tasolla oleva kuvailu hävittää joistakin aineistoista oleellisia asioita:

”Mut siellä on kanssa ihanat maisemat, jotka olis ollu vaan et jäämeri, ja eihän se riitä”

Tässä tapauksessa erityislaatuista videoaineistoa olisi kuvailtu asiasanalla, joka olisi johtanut siihen, että aineistosta olisi piilotettu sen oleellinen merkitys.

Haastatteluaineisto osoitti, että myös aineiston ajallisella kestolla on vaikutusta siihen, katoaako aineistosta jotakin merkittävää, kun sisällönkuvailu tehdään vain käyttäen asiansanoja. Tuotantokoordinaattori kertoi haastattelussaan, että uutiset saadaan kuvailtua käyttäen avainsanoja, koska uutisissa olevat yksittäiset uutisjutut kestävät vajaan kaksi minuuttia, toisin kuin hänen kuvailemansa dokumentit, jotka voivat olla tunnin mittaisia. Avainsanat voivat siis toimia paremmin, kun videoaineisto on lyhyt ja tiivis esitys jostakin aiheesta.

Erityyppiset aineistot siis vaativat erityyppisiä sisällönkuvailuja (taulukko 3). Siksi Ylellä on käytössä kaksi eri tapaa lähestyä sisällönkuvailuja. Aineistoja voidaan kuvailla käytössä olevien järjestelmän avulla kahdessa erityyppisessä käyttöliittymänäkymässä. Toinen näkymä pilkkoo ohjelman osiin ja näin ohjelma saadaan kuvailtua osa kerrallaan. Toinen taas mahdollistaa sisällönkuvailun koko ohjelmalle kerralla. Tästä huolimatta sisällönkuvailun ohjeistus on tuotantokoordinaattorin mukaan Ylessä rakennettu lähinnä palvelemaan uutisaineistoa.

”Sehän on niinkun uutistavaraan ajateltu enemmän, ja siinä mä sen ymmärrän, et jos sulla on puhuvia päitä ja joitain taloja Helsingissä ja jotain ulko/sisäkuva”

Uutisiakaan ei kuitenkaan kuvailla vain avainsanoja käyttäen, koska avainsanat eivät toisinedes uutisista oleellisinta asiaa esille. Avainsanoja käytetään uutisaineistossa täydentämään kuvailuja. Informaatikon haastattelusta ilmeni, että uutisten sisällönkuvailussa tärkeintä on kuvailtavan uutisjutun aihe. Aihe saadaan informaatikon mukaan vangittua sisällönkuvailuun helpoiten kirjoittamalla uutistenlukijan kolme ensimmäistä virkettä. Tämä on hyvä esimerkki aineistojen eroavaisuudesta sisällönkuvailun suhteen. Esimerkiksi Karjalasta kertovasta dokumentista tuskin saataisiin juuri mitään merkittävää sisältöä sisällönkuvailuun, jos kuvailuun kirjoitettaisiin dokumentissa esiintyvän henkilön kolme ensimmäistä virkettä.

Toisille aineistoille avainsanojen kirjoittaminen sisällönkuvailuun tuottaa riittävällä tasolla olevan kuvailun, eikä aineiston merkittävyys katoa esimerkiksi avainsanojen taakse.

Taulukko 3. Aineistot eivät ole homogeenisiä ja tarvitsevat siksi erityyppisiä kuvailuja

Sisältöjä pitää kuvailla monella tavalla jotta aineistojen kirjo tulee huomioitua	Jostain aineistoista menetetään jotakin merkittävää jos ne kuvaillaan liian yleisellä tasolla
	Erityyppiset videoaineistot tarvitsevat erilaista sisällönkuvailua
Aineistojen arvo vaihtelee jälleenkäyttöarvon ja merkittävyyden vuoksi	Arvokkaiden aineistojen sisällönkuvailu on historiallisesti merkittävää
	Toiset aineistot vaikuttavat enemmän sisällönkuvailun kehitykseen

5.4 Hyvin tehty sisällönkuvailu mahdollistaa jälleenkäytön

Tärkein syy sisällönkuvailujen tekemiseen on aineistojen jälleenkäyttöarvo (taulukko 4). Aineistoja kuvaillaan, jotta ne olisivat mahdollisimman hyvin löydettävissä tiedonhaussa. Videoaineistoja myös uusitaan, joten vanhat ohjelmat pitää pystyä löytämään. Jälleenkäyttöarvo on siis tiukasti sidoksissa tiedonhakuun (taulukko 4). Koska tiedonhaualla aineistojen jälleenkäyttö mahdollistetaan. Yksi syy aineistojen jälleenkäytölle on kustannustehokkuuden maksimoiminen. Ei ole järkevää viedä kuvauskalustoa sekä henkilöstöä metsään etsimään siiliä uutista varten, jossa kerrotaan siilien talvihorrostamisesta, vaan on järkevämpää etsiä arkistosta jo olemassa olevaa siilimateriaalia. Tämän vuoksi videoaineistojen kuvailussa oleellisinta onkin se mitä kuvassa näkyy. Tuotantokoordinaattori totesikin haastattelussa:

”Mähän en kirjoita sitä mitä ihmiset sanoo vaan sen on niinku kuvan määrittelemine, että mitä siinä kuvassa on”

Sisällönkuvailussa on otettava huomioon myös monenlaiset tiedonhakutilanteet ja niiden asettamat vaatimukset sisällönkuvailulle. Jos tiedonhaku on yleisellä tasolla, myös sisällönkuvailun tason on oltava yleinen. Vastaavasti hyvin tarkka tiedonhakutilanne vaatii yksityiskohtaisen sisällönkuvailun (taulukko 4).

Nykyisin videoaineistoja, aivan kuten muitakin aineistoja, haetaan tekstitiedonhakua hyödyntäen. Videoaineistot on tehty haettaviksi tekstitiedonhaulla kirjoittamalla niille tekstitiedonhaun ymmärtämä merkkimuotoinen sisällönkuvailu. Tiedonhaussa käytetään perinteisiä tiedonhaun menetelmiä. Järvelinin ja Sormusen (2000, 122) mukaan tiedonhakujärjestelmän oleellisin osa on täsmäytysalgoritmi. Täsmäytysalgoritmin tarkoituksena on laskea kyselyn ja dokumentin esityksen (representation) yhtäläisyys. Yhtäläisyyden perusteella hakujärjestelmä päättää, kuuluuko dokumentti tulosjoukkoon vai ei. Voidaan tulkita, että Järvelinin ja Sormusen tarkoittamaa esityksen samankaltaisuutta ei voi itse aineiston kanssa muodostua, jos aineisto ei ole tekstimuodossa. Tällöin täsmäytys tapahtuu aineiston kuvailun ja kyselyn samankaltaisuudesta. Tämä aiheuttaa ongelmia, koska ihmislähtöinenkin kuvailu on aina ihmisen tulkintaa kuvailtavana olevasta aineistosta.

Maisemakuvatyyppeiden epämääräisiksiin luokiteltavien tiedontarpeiden lisäksi tiedonhakijalla voi olla tarkka tiedontarve, joka asettaa lisävaatimuksia sisällönkuvailulle. Haastatteluissa ilmenneitä tarkkoja tiedontarpeita oli esimerkiksi kehityspäällikön mainitsema panssarivaunun tyyppi:

”Kun hänen piti löytää nimenomaan tietyn panssarivaunun kuva johonkin sotajuttuun”

Näin tarkkoja sisällönkuvailuja ei voida olettaa ihmisen tekevän. Sisällönkuvailun pitäisi siis vastata kahteen hyvin erityyppiseen tiedontarpeeseen, laveaan jopa tunnepohjaiseen tiedontarpeeseen ja hyvin yksityiskohtaiseen tiedontarpeeseen (taulukko 4).

Sen lisäksi, että erityyppiset tiedontarpeet vaativat erilaisia sisällönkuvailuja myös erityyppiset tiedonhakijat tarvitsevat erilaisia sisällönkuvailuja (taulukko 4). Esimerkiksi Yle Areenassa julkaistaan paljon Ylen materiaaleja. Vuonna 2016 Yle Areenassa on ollut tv-tallenteita 15 000 tuntia (Yle 2017). Julkaiseminen asettaa uusia vaatimuksia video-materiaalien sisällönkuvailulle, koska ennen Ylen materiaalit olivat tarkoitettu vain sisäiseen käyttöön ja niitä hakivat vain organisaatiossa työskentelevät henkilöt, jotka koulutettiin hakemaan aineistoja.

Markkula ja Sormunen (2006) ovat tutkineet Ylen sisäistä videotiedonhakua. He selvittivät, mitkä ovat tyypilliset työprosessin vaiheet, joissa videotiedonhakua tehdään, mitkä ovat videoaineistojen tyypillisimmät käyttökohteet ja millaiset videoaineistotarpeet kohdentuvat työprosessin eri vaiheisiin, sekä miten nämä tarpeet muuttuvat työprosessin edetessä. Tutkimuksessa videoaineistoja käyttävän tahon työprosessi jaettiin kuuteen eri vaiheeseen: ideointiin, suunnitteluun, kuvaamiseen, kuvien ensivalintaan, käsikirjoittamiseen ja kirjoittamiseen sekä editoimiseen. Markkulan ja Sormusen (2006) mukaan kaikissa näissä työprosessin vaiheissa suoritetaan omanlaistaan videotiedonhakua, kuitenkin pääosin videotiedonhaku tehdään prosessin kolmessa ensimmäisessä vaiheessa. Ideointivaiheessa tiedonhaku on taustatiedonhakutyypistä, suunnitteluvaiheessa tiedonhaku on informaation sekä materiaalien keräämistä ja kuvaamisen jälkeen haettiin arkistomateriaaleja lisättäväksi kuvattuihin kuviin. Pääosin Ylen sisäinen videomateriaalien tiedonhaku voidaan jakaa kahteen pääkategoriaan. Videoaineistoja käytetään informaatiolähteenä ja materiaalina. Aineistojen käyttäminen informaatiolähteenä tarkoittaa sitä, että aineistoja käytetään tukemaan uuden ohjelman ideointia, suunnittelua ja päätöksentekoprosessia sekä materiaalina, eli konkreettisina osina uutta ohjelmaa.

Tämä Markkulan ja Sormusen tutkimus kohdistui Ylen sisäisen videomateriaalien käyttöön, mutta uusia vaatimuksia videomateriaalien hakuun tuo nykyään se, että videomateriaalien hakijat eivät ole pelkästään Ylen sisäisiä. Kehityspäällikkö kertoi haastattelussa, että siinä on ero ovatko aineistojen jatkokäyttäjät talon sisältä vai yksityishenkilöitä. Kuvailu vastaa erilaisten tiedonhakijoiden tarpeisiin eri tavalla. Julkaiseminen on siis pakottanut Ylen arkiston pohtimaan, miten aineistoja sisältöjä kuvaillaan niin, että sisällönkuvailu palvelee myös hyvin toisenlaista kohderyhmää kuin ennen. Informaatikko totesikin haastattelussa metadatan tarpeen muutoksen tuovan uudenlaisia vaatimuksia sisällönkuvailuun.

Tiedonhakija on keskeisessä roolissa, koska hän luo kyselyn ja tiedonhakujärjestelmä hakee tekstiaineistosta tähän hakuun sopivia merkkijonoja. Videoaineistoissa merkkijonohaku on hankalaa, koska merkkijonomuodossa olevaa dataa on videoaineistossa verrattain vähän. Esimerkkejä näistä voisi olla vaikkapa teoksen nimi, tekstitystiedostot tai genre. Ongelmana ovat myös käyttäjän luomat kyselyt. Rautiaisen (2006, 23) mukaan tietokoneen suorittamaa täsmäytystä hankaloittaa käyttäjän tiedontarpeen artikuloinnin epäselvyys tehtäessä tiedonhakua, sekä epäselvyys siitä, mikä on aineiston kyky vastata

tähän tiedontarpeeseen. Epäselvyydet johtavat semanttiseen kuiluun käyttäjän ja järjestelmän välillä, joka vaikuttaa käyttäjän tyytyväisyyteen järjestelmää kohtaan. Rautiaisen näkemys on tulkittavissa niin, että hänen mukaansa tiedonhakija ei aina osaa tiedonhaku-järjestelmää käyttäessään ilmasta tiedontarpeitaan riittävällä tasolla. Tällöin aineisto ei voi vastata tiedonhakijan tiedontarpeeseen ja tämä aiheuttaa semanttisen kuilun. Esimerkiksi tiedonhakija voi etsiä videoaineistoa, joka sisältää hänen mielestään kaunista maisemakuva (taulukko 4). Epämääräinen tiedontarve, kuten kaunis maisemakuva, ei ole helposti siirrettävissä tekstimuodossa kyselyksi tietojärjestelmään. Tämän vuoksi aineiston kyky vastata kyseiseen tiedontarpeeseen on heikko, eikä tiedonhakija välttämättä saa tulokseksi hänen mielestään relevanttia tulosjoukkoa.

Semanttinen kuilu nousee siis ongelmaksi, kun tiedonhakijoiden tulisi löytää multimediallisista aineistoista relevantteja aineistoja. Järvelin ja Sormunen (2000, 118) määrittelevät relevanssin tiedonhaun kontekstissa seuraavasti: ”Relevanssilla tarkoitetaan informaation arvioitua käyttökelpoisuutta tietyssä käyttötilanteessa ottaen huomioon käyttäjän tavoitteet, arvot ja odotukset”. Videoaineistojen hakeminen ei ole aivan samanlaista kuin tekstitiedostojen hakeminen aineistojen multimediallisuuden vuoksi, eikä kuvailtava olevan objektin tärkein sisältö välttämättä löydy tekstiaineistoista. Näin ollen pelkästä tekstistä hakeminen aiheuttaa aineiston ja tiedonhakijan välille semanttisen kuilun, eikä käyttäjä löydä relevantteja aineistoja.

Taulukko 4. Jälleenkäyttö on mahdollista vain jos sisällönkuvailu on tehty hyvin

Videoaineistoista oleellista on tietää mitä kuvassa näkyy	Sisällönkuvailussa pitää kertoa mitä kuvassa näkyy
Erityyppiset tiedonhakutilanteet vaativat erityyppiset sisällönkuvailut	Yleisellä tasolla oleva tiedonhaku vaatii yleisen tason kuvailun
	Hyvin tarkka tiedonhaku vaatii tarkan sisällönkuvailun
Sisällönkuvailua tehdään jälleenkäytön vuoksi	Erityyppiset tiedonhakijat tarvitsevat erilaiset sisällönkuvailut

	Sisällönkuvailun toimivuutta arvioidaan tiedonhaun kautta
	Sisällönkuvailun tärkein motivaattori on jälleenkäytön mahdollistaminen

5.5 Videoaineistojen sisällönkuvailu ja aikadimensio

Videoaineistojen suhde aikaan tekee niistä erityisiä aineistoja. Ajallinen ulottuvuus vaikuttaa videossa olevien objektien relevanssiin, eli siihen miten kannattavaa objektia on kuvailla (taulukko 5). Objekti on relevantti kuvailun kannalta, jos sen mahdollinen relevanssi käyttäjälle on korkea. Ajallisella ulottuvuudella ja relevanssilla on siis suhde lähinä jälleenkäyttöarvon vuoksi. Kuvan, jossa kuvailtu objekti on, jälleenkäyttö on vähäisempi, mitä vähemmän aikaa objekti on näkyvillä. Huomioitavaa kuitenkin on, että jos objekti liittyy johonkin kokonaisuuteen, sen relevanssi kuvassa tämän tutkimuksen mukaan nousee.

Objektin roolin suuruus siis vaikuttaa objektin relevanssiin tiedonhaussa. Kehityspäällikön haastattelussa ilmeni, että jos objekti on kuvassa ”pääroolissa”, sen pitäisi tulla sisällönkuvailussa esiin (taulukko 5). Jos taas objekti on vain pieni kohde jossain kohtaa kuvaa, sen ei tarvitse välttämättä näkyä sisällönkuvailussa. Kiinnostavaa tässä on myös objektin paikka kuvassa ja sen vaikutus relevanssiin. Kehityspäällikön mukaan se, missä kohtaa kuvaa objekti on vaikuttaa myös relevanssiin:

”Et onks se keskellä onks se jossain taustalla olisi ehkä semmonen sen metatiedon ominaisuutena hyödyllinen asia”

Kehityspäällikön sanoista on tulkittavissa se, että ainakin sisällönkuvailusta tulisi selvitä, onko haettu objekti kuvassa pää- vai sivuroolissa. Tämän vuoksi objektin painoarvo kuvassa on oleellinen osa sisällönkuvailua. Myös muissa tutkimuksissa on tehty samankaltaisia havaintoja. Puhutaan siis semanttisesti rikkaasta kuvailusta, jossa esitellään objektit sekä niiden suhteet toisiinsa nähden. Smith (2001, 971) muistuttaa, että kuvan alueet liittyvät kuvan rakenteisiin, jotka sisältävät koostamisen, asettelun sekä yhteen liittämisen. Esimerkiksi kuvassa objekti voi olla suhteessa toiseen ”lähellä” tai ”vasemmalla”.

Ajallinen ulottuvuus on otettu huomioon Ylessä myös tuotantokoordinaattoreille tarkoitettussa sisällönkuvailun ohjeistuksessa. Tuotantokoordinaattori kertoo haastattelussa, että heitä on opastettu seuraavasti:

”7-10 sekuntia on semmonen minimi siihen mitä kannattaa lähteä erittelemään, paitti jos se on osa jotain kokonaisuutta”

Tätä ominaisuutta myös vaadittiin automaattiselta sisällönkuvailun ohjelmalta. Informaattikko kertoo, että hän odotti, että koneohjelma pystyisi tunnistamaan onko kuvajakso lyhyt vai pitkä. Automaattiselle sisällönkuvailun ohjelmalle pitäisi pystyä määrittelemään minkä pituisia kuvajaksoja kuvaillaan (taulukko 5). Olennaista tässä tuotantokoordinaatin kommentissa on kuvajakson pituuden lisäksi se, että kuvassa olevan objektin relevanssi kasvaa, jos se on osa jotakin kokonaisuutta.

Kuvien muodostama kokonaisuus vaikuttaa yksittäisen objektin relevanssiin. Jos yksittäinen objekti on kuvassa vain muutaman sekunnin, eikä se ole liitettävissä kokonaisuuteen, objekti katsotaan jälleenkäyttöarvoltaan niin alhaiseksi, että sitä ei kannata liittää sisällönkuvailuun. Automaattisen sisällönkuvailun ollessa kyseessä voidaankin esittää kehityspäällikön haastattelussa esiin tullut kysymys siitä, pitäisikö objektin sisällönkuvailuun liittää ominaisuutena objektin kuvassa oloaika. Yhden otoksen pituuden kuitenkin pitää vaikuttaa relevanssiin. Informaatikon mukaan:

”Sinällään jos haetaan kuvitusta johonkin kevättöihin, ni ne kaikki eri työvaihteet ni ne on käteviä, että ne on peräkkäin, se voi olla kokonaisuutena järkevä, vaikka näkyy yksittäisiä hirmulyhyitä pätkiä”

Objektin kuvassa oloaika voi siis olla lyhytkin ja silti relevantti, jos se on osa jotakin sellaista kokonaisuutta, joka koetaan merkitykselliseksi.

Kuvailun sitominen aikakoodiin on tärkeää erityisesti tiedonhakijalle (taulukko 5). Jos kuvassa näkyy erityisen hyvää lähikuvaa tärkeäksi koetusta henkilöstä, voidaan sisällönkuvailuun lisätä aikakoodin mukainen sisällönkuvailu ja kirjoittaa kuvassa olevan henkilön nimi. Esimerkiksi: "20:30:06 Tarja Halonen kävelee kadulla". Myös erityisen hienot maisemakuvat tai muut erityisenä pidettävät kuvat voidaan kirjoittaa sisällönkuvailuun ohjelmaan aikakoodin kanssa. Kuvasta 5 on nähtävissä aikakoodiin sidottu kuvailu. Kuvailu kertoo, mitä aihetta tämän aikakoodin sisällä käsitellään.



Kuva 5. Metro –järjestelmän aikakoodiin sidottu kuvailu.

Informaatikon mukaan tiedonhakijat haluavat aina päästä katsomaan suoraan ne kohdat mitä haettiin. Tiedonhakija ei halua kelata videoaineistoa siihen kohtaan, jossa häntä kiinnostava kohta aineistossa sijaitsee. Sen sijaan tiedonhakija olettaa, että video aukeaa hänelle relevantin aikakoodin kohdalta. Toisin kuin joissakin muissa sisällönkuvailun osuuksissa, aikakoodiin sidottu sisällönkuvailu ei tämän tutkimusaineiston mukaan vaihtelee riippuen siitä kuka on tiedonhakijana. Informaatikon mukaan julkaisuissa halutaan nähdä juuri se kohta videossa, missä tietty henkilö puhuu. Aikakoodiin sitomisessa toivottiin automaattiselta sisällönkuvailulta helpotusta. Toivomuksena oli, että automaattinen sisällönkuvailu osaisi esimerkiksi keskusteluiden perusteella arvioida, missä kohtaa videoaineistoa haettu henkilö puhuu. Tähän testissä ollut automaattisen sisällönkuvailun ohjelma pystyikin. Tunnistettuaan henkilön videoaineistosta automaattisen sisällönkuvailun ohjelma kykeni löytämään ne kohdat videosta, kun kyseinen henkilö puhui ja sitomaan aikakoodiin henkilön nimen. Näin kaikki aikakoodit, joissa henkilö puhui, oli helpposti löydettävissä.

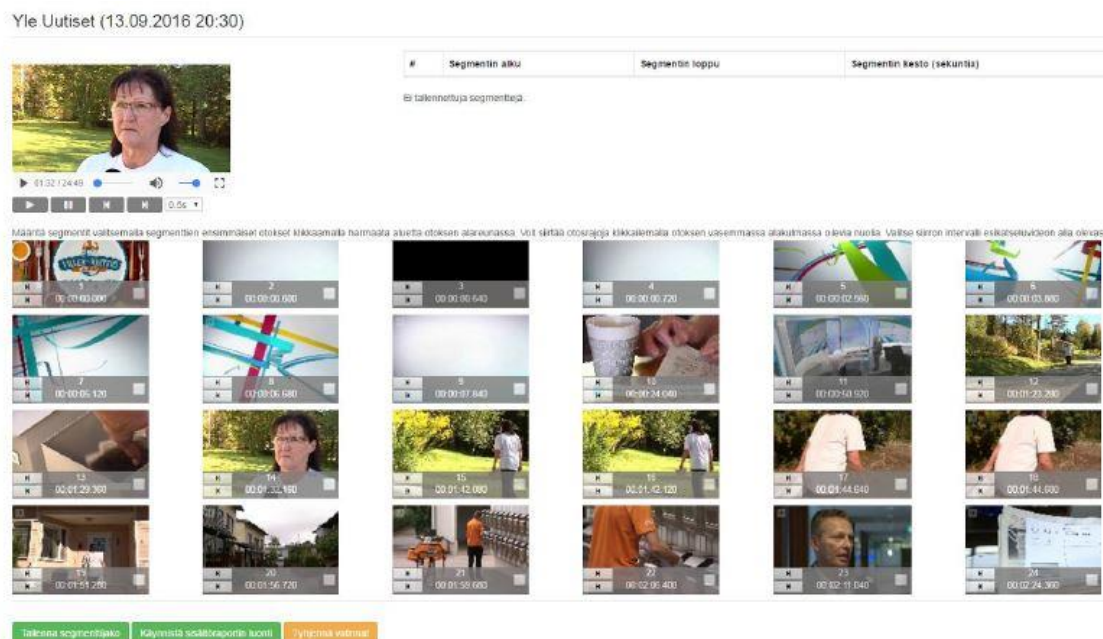
Sisällönkuvailu siis pitää sitoa aikakoodiin. Nykyisin puolen tunnin ohjelmat jaetaan tuotantokoordinaattorin mukaan noin viiden minuutin jaksoihin. Jokainen jakso kuvaillaan erikseen, mutta jaksoiden kuvailu voi olla myös sama. Informaatikon mukaan uutiset on helppo jakaa osiin, kun uusi jakso alkaa aina kun kuva palaa uutistenlukijaan.

Ohjelmien jakaminen osiin on olennainen osa sisällönkuvailua (taulukko 5). Jotta sisällönkuvailu voidaan suorittaa, täytyy ohjelmasta hahmottaa, mistä osista kokonaisuus koostuu. Nykyisin käytössä olevassa sisällönkuvailun ohjelmassa sisällönkuvailija näkee kuviksi pilkottuna melkein koko ohjelman kerralla (kuva 6).



Kuva 6. Metro –järjestelmässä sisällönkuvailijoiden on helppo hahmottaa ohjelman kokonaisuus.

Esimerkiksi informaatikko kertoi, että tehdessään sisällönkuvailua hän etenee rauhallisesti ohjelmaa segmentti kerrallaan. Kuten kuvassa numero kuusi on havaittavissa, Metro-järjestelmällä kuvan kokonaisuuden hahmottaa helposti. Kuvasta numero voidaan päätellä, että automaattisen sisällönkuvailun ohjelmalla kuvailtavan aineiston kokonaisuuden hahmottaminen ei ole aivan yhtä helppoa (kuva 6). Informaatikko toivoi, että automaattisen sisällönkuvailun ohjelma olisi tehnyt valmiiksi jonkinlaisen ehdotuksen siitä, mihin segmenttirajat ohjelmassa sijoittuisi. Sen sijaan automaattisen sisällönkuvailun ohjelma tuntui tarjoavan vain satunnaisia kuvia segmenttirajoiksi. Kuvassa 7 on automaattisen sisällönkuvailun ohjelman ehdotelma Yle Uutisten segmenttirajoiksi. Kuten kuvasta voidaan päätellä, kuvat näyttäytyvät satunnaisina käyttäjälle ja käyttäjän tuli itse valita sopivat kohdat segmenteille selailemalla kuvia.



Kuva 7. Automaattisen sisällönkuvailun ohjelma ehdottaa segmenttirajoille paikkoja satunnaisilla kuvilla.

Jotta automaattisen sisällönkuvailun ohjelma olisi voinut jakaa ohjelman osiin, sen pitäisi pystyä medioiden avulla hahmottamaan ohjelman ajallista dimensiota (taulukko 5). Kehityspäällikkö totesi, että automaattinen sisällönkuvailu voisi esimerkiksi aiheanalyysin tai tekstityksen pohjalta jakaa ohjelman osiin. Automaattisen sisällönkuvailun ohjelma voisi kehityspäällikön mukaan tehdä oletuksen, että jos aihe ohjelmassa, esimerkiksi uutisissa muuttuu, todennäköisesti myös yksi ohjelman osuuksista, kuten yksittäinen uutinen loppuu.

Videoaineiston kokonaisuuden hahmottamisen tärkeydestä kertoo puolestaan se, että tuotantokoordinaattori mainitsi muutamaaan kertaan, kuinka vaikeaa testattavana olleen automaattisen sisällönkuvailun ohjelman käyttöliittymältä oli hahmottaa ohjelmaa kokonaisuutena. Hän koki, että tämä kokonaisuuden hahmottamisen vaikeus hankaloitti osaltaan ohjelman kuvailemista automaattisen sisällönkuvailun ohjelmalla.

Taulukko 5. Videoaineistojen kuvailussa aika on otettava huomioon

Kuinka kauan ja suuresti objekti on kuvassa, vaikuttaa sen relevanssiin	Objektin roolin suuruus vaikuttaa sen relevanssiin
---	--

	Kuinka kauan objekti on kuvassa vaikuttaa kohteen kuvauksen relevanssiin
	Yhden otoksen pituuden pitää vaikuttaa relevanssiin
Tiedonhakijalle on tärkeää, että kuvailu on sidottu aikaan	Tiedonhakijana haluan päästä suoraan katsomaan oikeaa hakemaani videosta
	Kuvailu pitää sitoa aikaan (aikakoodiin)
Jotta voin kuvailla, minun pitää hahmottaa ohjelman kokonaisuus ja sen osat	Koneen pitäisi pystyä medioiden avulla hahmottamaan ajallista dimensiota
	Ohjelman kokonaisuuden hahmottaminen on tärkeää
	Ohjelman osioiden tunnistaminen on tärkeää
	Ohjelman jakaminen osiin on tärkeä osa sisällönkuvaailua
	Koneen pitäisi pystyä medioiden avulla hahmottamaan ajallista dimensiota
	Ohjelman kokonaisuuden hahmottaminen on tärkeää
	Ohjelman osioiden tunnistaminen on tärkeää
	Ohjelman jakaminen osiin on tärkeä osa sisällönkuvaailua

5.6 Sisällönkuvailu vaatii tietojen yhdistämistä

Samaan tapaan kuin ihmisenkin, on koneen käytävä läpi oppimisprosessi, jotta se voi tunnistaa objektit oikein (taulukko 6). Siksi koneelle ja ihmiselle yksityiskohtainen kuvailu on vaikeaa, koska se vaatii asian erityistuntemusta. Yksityiskohtaista tietoa tarvitaan esimerkiksi ihmisten tunnistamiseen ja nimeämiseen. Automaattisen sisällönkuvailun taustalla toiminut tekoäly kykeni yhdistämään henkilön nimen ja kasvot opittuaan asian. Informaatikko kertoi:

”Heidän sinne tietokantaansa se tallentuu tällöisiä kasvoja ja siitä sen kasvon eri ilmentymiä kun se käyttäjä vahvistaa, että kyllä tämä on Sauli Niinistö meidän presidenttimme”

Tässä tapauksessa automaattisen sisällönkuvailun ohjelma tallensi tietokantaansa saman henkilön kasvojen ilmentymiä. Kun käyttäjä vahvistaa, että kyseessä on tietty henkilö, esimerkiksi Sauli Niinistö, yhdistää automaattisen sisällönkuvailun ohjelma kaikki kasvoilmentymät tähän nimeen.

Tekoälyn oppiminen ei kuitenkaan rajautunut vain henkilöihin, vaan koneen tuli oppia tunnistamaan kaikki objektit. Oletus siis on, että ihmisen tulee myös opettaa koneelle kaikki kuvassa näkyvät objektit, jotta ne olisi mahdollista sisällönkuvailussa myöhemmin tunnistaa. Oppiminen ei myöskään tapahdu yhdellä kertaa, vaan vaatii useamman toiston. Yksityiskohtainen sisällönkuvailu vaatii kuitenkin myös ihmiseltä erikoistietämystä (taulukko 6). Näin ollen konetta opettavan ihmisen tulee tietää mitä kuvassa on. Esimerkiksi paikkojen tunnistaminen kuvan perusteella on myös ihmiskuvailijalle vaativaa. Haastattelusta selviää, ettei kuvituskuvaksi valikoituneen kaupunkikuvan paikan tunnistaminen ole helppoa. Videoaineistosta ei välttämättä käy lainkaan ilmi se, missä kohta on kuvattu. Tällöin vain henkilö, joka on ollut mukana tuotantoprosessissa, omaa riittävät tiedot kuvan täydelliseen kuvailuun.

Se, että kone tulee opettaa tunnistamaan objekteja ihmisvoimin sekä sisällönkuvailijoille asetetut vaatimuksen yksityiskohtaisen sisällön tuottamiseksi johtavat siihen, että kone ei tuota riittävästi yksityiskohtaista tietoa. Tämän tutkimuksen kohteena olleessa automaattisen sisällönkuvailun testissä kone pystyi esimerkiksi tunnistamaan lipputangon ja lipun, muttei sitä mikä lippu on kyseessä. On kuitenkin huomioitava, että kone olisi todennä-

köisesti pystynyt kertomaan, mikä lippu lipputangossa oli, jos se olisi aiemmin sille opetettu. Kehityspäällikön mukaan testauksessa mukana olleet henkilöt eivät päässeet yksimielisyyteen siitä oliko koneen tuottama sisällönkuvailu liian yksityiskohtaista vai ei:

”Niitä kuvapinnan tulkintoja, niitä tägityksiä mitä esineitä siellä on, niin osa piti just ihanteellisen yksityiskohtaisena ja osa piti liian yksityiskohtaisena.”

Testissä ollut automaattisen sisällönkuvailun ohjelma pystyi esimerkiksi tunnistamaan lähes poikkeuksetta kuvassa näkyvät solmiot. Osa testinkoehenkilöistä tuntui olevan sitä mieltä, että jatkuva solmioiden tunnistaminen oli turhaa sekä lisäsi sisällönkuvailuun turhaa hälyä. Osa taas oli sitä mieltä, että esimerkiksi tehtäessä dokumenttia vaikkapa miesten solmiomuodin muuttumisesta 2000-luvulla olisi solmioiden tunnistaminen erittäin hyödyllistä sisällönkuvailua.

Eräs sisällönkuvailun tärkeimmistä elementeistä on tietojen tehokas erottelu ja yhdisteleminen (taulukko 6). Automaattisen sisällönkuvailun ohjelman pitäisikin siis osata yhdistää monen eri tietolähteen tiedot, jotta se voisi tuottaa semanttisesti rikasta ja laadukasta sisällönkuvailua. Kehityspäällikkö huomautti, että usein oleellinen tieto on vain josakin multimediallisen aineiston osassa, kuten tekstityksessä tai kuvassa. Jos taas rinnakkain olevissa eri tiedonlähteissä kaikki indikoivat samaa asiaa, on tämä vahvempi todiste tiedon oikeellisuudesta. Jos videoaineiston haastattelussa puhutaan vaikkapa koirien kouluttamisesta, planssitekstissä on luettavissa haastateltavan olevan ammatiltaan koirankouluttaja ja taustalla on näkyvissä koiria, voidaan olettaa, että silloin hyvin todennäköisesti kyseessä on koirankouluttamiseen liittyvä aineisto.

Naphide ja Huang (2001) huomauttavat, että semanttiset konseptit ovat yhteydessä toisiinsa. Joidenkin avainsanojen tunnistaminen korottaa todennäköisyyttä sille, että toiset avainsanat ovat oikeellisia. Jos ohjelma tunnistaa taivasta ja vettä, nostaa se avainsanan ranta todennäköisyyttä olla oikein ja samalla vähentää avainsanan sisätilat todennäköisyyttä olla oikein. Tietojen yhdistämisen kautta tapahtuva sisällönkuvailun oikeellisuuden arvioinnin ei kuitenkaan pitäisi rajoittua vain kuvantunnistukseen, vaan tietoja pitäisi voida tunnistaa useasta eri mediasta.

Useasta mediasta saadun tiedon yhdistelemisen lisäksi myös aiemmin tuotetun datan hyödyntäminen on oleellista sisällönkuvailun prosessin tehostamiseksi (taulukko 6). Järjestelmien integroituminen helpottaisi siten, ettei jo kertaalleen syötettyä dataa olisi tarvetta

syöttää järjestelmään enää uudelleen. Usein osa tiedoista syötetään järjestelmiin tuotannon eri vaiheissa. Näiden tietojen käyttöönotto olisi tehokasta, oli kyseessä sitten koneen tai ihmisen tekemä sisällönkuvailu. Tehokas automaattisen sisällönkuvailun ohjelma tekisi kuitenkin tietoja yhdistävää työtä itsenäisesti.

Eri tietoja pitäisi myös pystyä erottelemaan toisistaan (taulukko 6). Informaatikon mukaan tietojen erottelu erillisiin kenttiin on osa heidän tietomalliaan. Tietojen erottamisen lisäksi eri medioista saatua informaatiota pitäisi pystyä myös yhdistämään kuvasta saatuun informaatioon. Esimerkiksi tekstitystiedostosta saatua informaatiota pitäisi pystyä yhdistämään kuvaan. Testissä ollut automaattisen sisällönkuvailun ohjelma ei kyennyt tarpeeksi tehokkaaseen tietojen yhdistämiseen. Informaatikko kertoi keskustellessaan ohjelmaa kehittäneiden henkilöiden kanssa siitä, että esimerkiksi tekstin ja kuvan yhdistämisen tulisi ohjelmassa toimia paremmin.

Testissä olleessa automaattisen sisällönkuvailun ohjelmassa ei ollut ominaisuutena puheentunnistusta. Tämä koettiin testauksessa ongelmalliseksi. Puheentunnistuksen puuttuminen saattoi kehityspäällikön mukaan aiheuttaa sisällönkuvailuun laadullisia ongelmia:

”Se relevanssin arviointi audion kautta varmasti toisi lisää työkaluja laadun parantamiseen”

Puheentunnistuksen puuttumisen koettiin yleisesti vaikuttavan erityisesti sisällönkuvailun sisällön relevanssin arviointiin (taulukko 6). Jos kuvantunnistuksessa tunnistetut objektit löytyvät myös sanoina puheentunnistuksen avulla, voitaisiin päätellä, että objekti on relevantti liitettäväksi sisällönkuvailuun. Näin kuvien kuvailu saisi kontekstia äänen avulla. Rautiaisen (2006, 29) mukaan audiolla on äärimmäisen tärkeä rooli videoaineistojen sisällön ymmärtämisessä. Tämän perusteella voidaan ajatella, että audiolla on vaikutusta aineistojen koko semantiikan saavuttamiseen. Aineistoissa kuten elokuvat, mainokset tai musiikkivideot ääni välittää tärkeää informaatiota aineiston sisällöstä. Esimerkiksi uutisaineistojen automaattisessa kuvailussa puheentunnistuksen voidaan olettaa olevan tärkeää. Uutisissa on hyvin vähän kuvia, joiden kautta pelkällä kuvantunnistuksella päästäisiin mielekkääseen lopputulokseen. Uutisissa luetaan esimerkiksi sähkötyyppejä uutisia, joista ei kerrota kuvalla välttämättä mitään. Pelkkää kuvantunnistusta käyttäen näistä otteista ei jää pois pelkästään semantiikka vaan koko uutinen. Multimediallisissa aineistossa on paljon sisältöä, jonka kokonaisvaltaista tulkintaa tarvitaan, että saadaan vangittua aineiston koko semantiikka.

Ylellä on aikaisemmin tutkittu puheentunnistuksen toimivuutta toimittajien tekstien litteroinnissa. Puheentunnistuksen perusteella toimivan litteroinnin koettiin toimivan riittävän hyvin. Vaikka puheentunnistus ei ollut absoluuttisen tarkkaa, valmis teksti kuitenkin helpotti toimittajien työtä heidän litteroidessaan haastatteluita. Tuotantokoordinaattori luonnehtii asiaa näin:

”Vaikka ne oli huonoja, ni yks mun toimittaja sano että seki auttaa, et ei sen tarte 100 prosentista olla, kuhan saa niinku jonku liuskan ni kyllä hän muistaa. Et se hänen mukaansa tosi paljon auttoi. Hän säästi monta tuntia työtä sillä.”

Tässä on kuitenkin otettava huomioon litteroinnin ja sisällönkuvailun erot. Vaikka litteroinnissa toimittaja pystyi hyväksymään tekstin, joka oli vain osiltaan tarkkaa, sisällönkuvailija voi kokea samalla tarkkuudella tehdyn sisällönkuvailun huonoksi ja hyödyttömäksi. Tämä voi johtua siitä, että sisällönkuvailija on tottunut työssään ihmisen tuottamaan laadukkaampaan sisällönkuvailuun.

Puheentunnistuksen puuttuminen kuitenkin jätti kysymyksen siitä olisiko automaattisen sisällönkuvailun ohjelman tuottama sisällönkuvailu ollut laadukkaampaa, jos siihen olisi vielä yhdistetty puheentunnistuksesta saatu informaatio. Informaatikko haastattelussaan suosittelee automaattisen sisällönkuvailun ohjelman tekijöille tiiviimpää yhteistyötä puheentunnistusta tuottavien yritysten kanssa. Eikä informaatikko ollut mielipiteensä kanssa yksin. Ylen automaattisen testauksen loppuraportissa Saarikoski (2016) kirjoitti, että puheentunnistuksen käyttäminen henkilöiden tai videokuvan tunnistamisen rinnalla olisi ollut kannattavaa.

Automaattisen sisällönkuvailun taustalla toimi ontologia. Ontologia ei kuitenkaan ollut käyttäjille näkyvä ja tämä koettiin ongelmaksi (taulukko 6). Varsinkin siksi, että ontologian laatu vaikuttaa koko sisällönkuvailun laatuun. Ontologian rakenteen tietäminen olisi ulkopuolisen arvioijan tai koekäyttäjän näkökulmasta merkittävä kriteeri. Kehityspäällikön mukaan kuitenkin on harvinaista, että kerrotaisiin julkisesti mikä ontologia tai taksonomia on käytössä. Arvioijille olisi tärkeää hänen mukaansa myös tietää, minkälaiseen rakenteeseen mikäkin palvelu tukeutuu.

Osa sisällönkuvailijoista käyttää sisällönkuvailua tehdessään apunaan tekstitystiedostoja. Tuotantokoordinaattori kertoi, että sisällönkuvailua tehdessään hän laittaa aina tekstitystiedoston näkyville. Tuotantokoordinaattori täsmensi, että hänen tarpeensa tekstitystie-

dostojen käyttämiseen sisällönkuvailun apuna johtuu siitä, että suomi ei ole hänen äidinkieltään. Tämän vuoksi hän joutuu hakemaan sanoja tekstitystiedostosta. Vähän samaan tapaan toimii myös automaattisen sisällönkuvailun ohjelma. Sekä tuotantokoordinaattorilla että automaattisen sisällönkuvailun ohjelmalla, ongelmia esiintyy sisällönkuvailussa, kun tekstitystiedosto puuttuu. Kehityspäällikkö muistutti, että jos sisällönkuvailussa tukeudutaan liikaa tekstitystiedostoihin, niin se kaventaa ohjelman soveltuvuutta kaikilta niiltä ohjelmilta, missä ei ole tekstitystiedostoa tarjolla.

Kuitenkin jos tekstitystiedostoja tai muita tekstitiedostoja on tarjolla, kone tekee parempaa sisällönkuvailua (taulukko 6). Kun kone tunnistaa kuvapinnasta objektin se voi hakea tekstitystiedostosta vahvistuksen tunnistukselleen. Esimerkiksi jos automaattisen sisällönkuvailun ohjelma tunnistaa henkilön, joka voi olla presidentti Sauli Niinistö ja tekstitystiedostosta löytyy merkkijonot Sauli ja Niinistö kasvattaa se sisällönkuvailun oikeellisuuden todennäköisyyttä. Automaattisen sisällönkuvailun ohjelma pystyi myös analysoimaan ohjelman aihetta tekstitystiedostojen avulla, vaikka toistaiseksi se ei pystynytään yhdistämään aihetta kuvaan. Aiheiden poimiminen tekstitystiedostoista onnistuu kuitenkin informaatikon mukaan kohtalaisen hyvin. Kuten edellä mainittiin, tekstitystiedostojen hyödyntämisen suurin ongelma on se, etteivät kaikki aineistot sisällä minkäänlaista tekstitystiedostoa.

Taulukko 6. Sisällönkuvailu on mahdollista vain yhdistämällä taustatietoja ja eri medioista kerättyä dataa

Koneelle ja ihmiselle yksityiskohtainen kuvailu on vaikeaa	Ihan niin kuin ihmisenkin, koneenkin pitää oppia tunnistamaan asioita
	Yksityiskohtainen sisällönkuvailu vaatii ihmiseltäkin erikoistietämystä
	Kone ei tuota riittävästi yksityiskohtaista tietoa
Tehokas järjestelmä osaa erotella ja yhdistää tietoja	Automaattisen sisällönkuvailun pitäisi osata yhdistää monen eri tietolähteen tiedot

	Tehokas sisällönkuvailun ohjelma hyödyntää myös aiemmin tuotettua dataa
	Eri tietoja pitää pystyä erittelemään toisistaan
	Automaattisen sisällönkuvailun pitäisi osata yhdistää monista eri medioista saadut tiedot
Puheentunnistus voisi auttaa automaattisessa sisällönkuvailussa	Puheentunnistus voisi auttaa relevanssin arvioinnissa
	Meillä on kokemusta siitä, että puheentunnistus helpottaa työtä
Tekstitiedostot ovat oleellinen osa sisällönkuvailua	Tekstitystiedoston puuttuminen vaikeuttaa sisällönkuvailua
	Tekstitystiedostojen avulla kone tuottaa parempaa sisällönkuvailua

6 MIKSI KONE EI RIITÄ SISÄLLÖNKUVAILUSSA

Tässä luvussa tarkastellaan empiirisiä tuloksia, jotka osoittavat miksi automaattinen sisällönkuvailun ohjelma ei vielä riitä korvaamaan ihmislähtöistä sisällönkuvailua. Luvussa keskitytään erityisesti siihen, mitä lisäarvoa ihminen tuo sisällönkuvailuun.

6.1 Ihminen kertoo sen mitä kuvassa ei näy

Hyvä sisällönkuvailu kertoo ohjelmasta jotain, mikä ei varsinaisesti näy kuvassa. Videoaineistolla on yleensä jokin aihe, asia jota käsitellään. Uutisaineistojen tapauksessa aihe voi olla helposti löydettävissä. Uutistenlukija tiivistää useimmiten uutisen aiheen kolmessa ensimmäisessä virkkeessä. Tämä olisi informaattikon mukaan automaattiselle sisällönkuvailun ohjelmalle helposti opetettavissa. Jos automaattinen sisällönkuvailun ohjelma osaisi poimia uutistenlukijan kolme ensimmäistä virkettä suoraan sisällönkuvailuun, tämä helpottaisi merkittävästi sisällönkuvailijan työtä. Toinen asia on se, pystyisikö automaattisen sisällönkuvailun ohjelma tuottamaan informaattikon vaatimaa näkökulmaa aiheeseen.

”Et sen lisäksi me halutaan se suoranainen kertomus siitä aiheesta. Sen lisäksi, että puhutaan että aiheena on Ahvenanmaan puolustus ja näin, et me halutaan myös et mikä se näkökulma tähän aiheeseen on”

Abstraktimpaa semanttista tietoa lisätään sisällönkuvailuun esimerkiksi avainsanojen tai asiasanojen avulla (taulukko 7). Informaatikko pohtii, mikä on asiasanoituksen ja suoranaisen aiheen ero. Hänen mukaansa erovaisuus on siinä, että asiasanoituksella ja aiheella on erilainen painotus. Aiheena voi olla esimerkiksi se, että tänä vuonna on lämpimämpää kuin viime vuonna. Asiasana edelliseen aiheeseen kuitenkin olisi ”ilmastonmuutos”. Informaatikko antoi myös toisen esimerkin aiheen ja asiasanoituksen erosta:

”Puhutaan vaikka Hornet-kaupoista niin sitten siihen lisää siihen tällöisiä niinku puolustusvoimat”

Ihminen siis pyrkii tuottamaan sisällönkuvailuun yleisemmän tason (taulukko 7). Tällä yleisemmällä tasolla pyritään tuottamaan lisäarvoa tiedonhakijalle. Oletetaan, että kuvailussa olisi pelkästään kuvasta tunnistettuja objekteja. Objektit voisivat olla esimerkiksi

”lehmä” ja ”traktori”. Tiedonhakijan halutessa kuvituskuva maalaismaisemasta, hänen tulisi osata määritellä hyvin tarkkaan minkä tyyppisiä objekteja liittyy maalaismaiseen. Jos hän hakee esimerkiksi sanoilla ”lato” ja ”laidun”, haku ei täsmää edellä kuvailtuun videoaineistoon ja aineisto jää pois hakutulospöytäkirjasta. Sisällönkuvailija saattaisi lisätä kuvailuun asiasanan, kuten maatila tai maatilan pihapiiri, joka nostaisi kuvailun tasoa ylemmäs ja tiedonhakijan ei tarvitsisi määritellä yksittäisiä objekteja, joita haluaa kuvassa näkyvän. Informaatikon mielestä ainakaan testissä ollut Valossan tekemä automaattisen sisällönkuvailun ohjelma ei tuottanut ylemmän tason sisällönkuvailua.

Yksi tapa tuoda kuvailuun ylempää tasoa, on tukeutua taustatietoihin. Esimerkiksi Dalakleidi ja kumppanit (2011, 25) korostavat taustatietojen käytön tärkeyttä, jos ongelman ratkomisen puhtaasti data-lähtöisesti johtaa puutteelliseen lopputulokseen. Tämä nousee esille etenkin tapauksissa, jossa lopputulos olisi erilainen, jos siinä olisi sisällönkuvaajan tekemä havainto siitä, mitä tietoa haluttiin välittää. Videoaineistojen taustatietoina voitaisiin pitää teoksen nimeä, tekijää tai vaikkapa julkaisuvuotta. Semanttisesti rikasta sisältöä voitaisiin siis saada käyttöön, jos automaattinen sisällönkuvailu osaisi ottaa huomioon kontekstin, jossa aineisto on luotu. Kontekstilla tässä yhteydessä tarkoitetaan esimerkiksi laajempaa kulttuurista tai yhteyskunnallista viitekehystä. Joitakin videoaineistoja kuvailtaessa voisi olla esimerkiksi tärkeää tietää, minä historian ajankohtana video on luotu. Tämä tarkoittaisi käytännössä sitä, että automaattisen sisällönkuvailun tulisi ottaa huomioon sellaisiakin asioita, jotka eivät varsinaisesti näy kuvassa. Enser (1999) antaa tällaisesti tapauksesta hyvän esimerkin. Lentokoneen kuvalla on hyvin vähän merkitystä (paitsi kuvaajalle itselle) ilman kuvaa selittävää selostetta, josta selvää, että lentokoneen kyydissä istuu Chilen entinen presidentti Augusto Pinochet. Kuvassa oleva lentokone on kuljettamassa Pinochetia vapauteen. Tämä näkymätön fakta paljastaa sen, että kuvalla on suuri metaforinen merkitys.

Samankaltaisia esimerkkejä voidaan löytää Ylen kontekstista. Voitaisiin esimerkiksi hakea Ylen videoaineistoja, joka kertovat miten Suomen poliittinen tilanne muuttui uuden presidentin myötä. Tiedontarve on melko abstrakti eikä aineistossa, joka tähän tiedontarpeeseen olisi relevantti, välttämättä mainita lainkaan sanaa presidentti. Enserin (2008, 538) mukaan rajattu objektien ja kohtausten havaitseminen aiheuttaa ongelmia, koska ohjelmat ovat opetettu tunnistamaan vain niitä ominaisuuksia, jotka ovat kuvassa näkyvillä. Tämä viittaa siihen, että varsinkin multimediallisten aineistojen semanttisuus ulot-

tuu paljon syvemmälle kuin vain siihen, mitä kuvassa varsinaisesti näkyy, mitä äänimaa-ilmassa kuuluu tai mitä tekstiaineistossa kerrotaan. Multimediallisten aineistojen semanttisuus ulottuu ylitse aineiston omien rajojen ja sen ymmärtämiseen tarvitaan ainakin vielä ihmistä.

Jos tekstitiedonhaku, joka on ulotettu hakemaan myös taustatiedoista, ratkaisisi videotiedonhaun ongelman, Ylellä ei olisi enää tarvetta kehittää automaattisen sisällönkuvailun ohjelmaa. Ylellä videosisältöjen hakeminen on Viljasen ja Mattilan (2017) mukaan riippuvaista niihin liitetystä tekstistä ja muusta metatiedosta, kuten otsikoista, kuvausteksteistä, genreistä ja kansikuvasta. Riippuvaisuus edellä mainituista metatiedoista on kuitenkin ongelmallista, koska multimedialliset aineistot sisältävät tekstin lisäksi esimerkiksi kuvaa, musiikkia ja puhetta. Tällöin tiedonhaun ulottaminen taustatietoihin ei ratkaise ongelmaa, joka Ylellä on tunnistettu. Kuitenkin voidaan esittää oletus siitä, että taustatietojen avulla kontekstin liittäminen automaattiseen sisällönkuvailuun voisi nostaa kuvailun yleisemmälle tasolle ja näin parantaa kuvailun semantiikkaa.

Yleisemmän tason lisäksi ihminen tuo sisällönkuvailuun kontekstin avulla semantiikkaa (taulukko 7). Esimerkiksi, jos kuvassa näkyy traktori noin viisi sekuntia ja tämän jälkeen kuva siirtyy peltomaisemaan, haastattelussa ollut informaatikko lisäisi sisällönkuvailuun sanan ”auraaminen”. Vaikka Enserin (1999) mukaan kuvassa olevan merkityksen vangitseminen on älyllisesti vaativa tehtävä (Enser, 1999), tässä tapauksessa sisällönkuvailija ymmärtää, mikä merkitys traktorilla on pellolle, joten hän kontekstin avulla nostaa sisällönkuvailun semantiikkaa. Tämän vuoksi konteksti on sisällönkuvailussa välttämätöntä. Jos tiedonhakija haluaa esimerkiksi videomateriaalia presidentti Sauli Niinistöstä, tiedonhaku täsmää kaikkiin niihin henkilöihin, jotka ovat presidenttimme kaimoja. Kontekstin avulla tiedonhakija voi päätellä, onko kyseessä tasavallan presidentti Sauli Niinistö vai joku toinen henkilö.

Taulukko 7. Ihminen kertoo sen mitä kuvassa ei näy, mutta mitä se on ja mihin se liittyy (konteksti)


Ihminen tarvitaan tuottamaan kuvailuun se mitä kuvassa ei näy	Aihe antaa pohjan sisällönkuvailulle
	Sisällönkuvailija nostaa kuvailun tasoa ylemmäksi asiasanojen avulla

	Ihminen pyrkii tuottamaan kuvailuun tarvittavan yleisemmän tason
Ilman kontekstia kuvailusta jää puuttumaan oleellista tietoa	Ihminen tuo sisällönkuvailuun kontekstin avulla semantiikkaa
	Konteksti on sisällönkuvailussa välttämätöntä

6.2 Automaattisen sisällönkuvailun oppimiskyky on rajallinen

Automaattisen sisällönkuvailun ohjelman oppimien on sidottu ihmisiin. Tämä johtuu siitä, että automaattinen sisällönkuvailu vaatii jatkuvat ihmislähtöistä opettamista (taulukko 8). Käytännössä tämä tapahtuu niin, että kone tunnistaa kuvapinnasta objektin ja antaa käyttäjälle ehdotuksen siitä mikä objekti voisi olla. Kuvassa 8 on nähtävissä millaisessa muodossa käyttäjälle aihe-ehdotukset käyttöliittymällä esiteltiin. Käyttäjä hyväksyi automaattisen sisällönkuvailun avainsanaehdotuksen painamalla oikeinmerkkiä tai hylkäsi ehdotuksen painamalla raksia. Käyttäjän vahvistaessa ehdotuksen ohjelma sai tiedon siitä, että tunnistus osui oikeaan. Vaihtoehtoisesti käyttäjä pystyi myös korjaamaan ehdotuksen, jos se on hänen mielestään ollut virheellinen.

A-studio: Talk (20.10.2016)



The screenshot shows the A-studio interface. On the left is a video player with a woman speaking. Below the video are playback controls (play, pause, stop, next, previous, volume, and a 0.5s speed selector). Below the video player, it says "Segmentti 2 (00:00:14.200)". Below that are two tabs: "Yhteenveto" and "Henkilöt". At the bottom, there is a section titled "Aiheet" (Topics) with a grid of topic tags. Each tag has a checkmark icon and a small 'x' icon. The tags include: demilitarisointi, ahvenanmaalainen, intressi, turvallisuustilanne, puolustusministeri, kannanotto, demilitarisoitu, sotilaallinen, merili, Jukka Tarkka, Iso-Britannia, Pohjois-Itämeri, Soini, Geneve, Isoa-Britannia, Neuvostoliitto, Helsinki, Ahvenanmaa, Maarianha, Tukholma, Suomi, Pariisi, Ukraina, Oolanti, Krim, Itämeri, Venäjä.

Kuva 8. Automaattinen sisällönkuvailun ohjelma ehdottaa A-studioon sopivia aiheita.

Kehityspäällikön mukaan automaattisen sisällönkuvailun ohjelma vaatii jatkuvaa opettamista ainakin alkuvaiheessa ja myöhemmin jonkinlaisella syklillä. Vastuu ohjelman opettamisesta mietitytti kehityspäällikköä. Onko esimerkiksi uusien termien tai uusien ilmiöiden opettaminen vain Ylen omalla vastuulla? Automaattisen sisällönkuvailun ohjelmaa opettavan tahon täytyy myös kantaa vastuuta siitä, mitä ohjelmalle opetetaan. Miten kone esimerkiksi ottaa vastaan uusia ärsyksiä ja miten se näihin ärsyksiin reagoi? Tähän liittyy myös eettisiä kysymyksiä, siitä onko sisällönkuvailun määrittäminen Ylen omissa käsissä vai tekeekö opettamisen ylläpitävä taho. Automaattisen sisällönkuvailun ohjelman opettaminen vaatii paljon työtä. Työ ei ole ainoastaan konkreettista, jossa korjataan sisällönkuvailun virheitä, vaan myös määrittelytyötä vastuista. Näiden ongelmien pohtiminen johtuu osaltaan siitä, että videoaineistojen automaattinen sisällönkuvailu on vielä sellaisessa vaiheessa, ettei valmista automaattisen sisällönkuvailun ohjelmaa vielä ole olemassa. Informaatikon mielestä mahdollisuuksia on kyllä paljon, mutta käytännön sovellukset eivät ole vielä aivan käyttäjien sormien ulottuvilla. Kehityspäällikön sanoja lainaten:

”Ei oo semmonen niinkun kertalaakista käynnistettävä sisällönkuvailun automaatti”

Konetta pitää vielä opettaa tekemään laadukasta sisällönkuvailua. Tämä opettaminen ei kuitenkaan tämän tutkimuksen mukaan ollut aivan niin helppoa, kuin sen olisi toivottu olevan. Ongelma automaattisen sisällönkuvailun ohjelman testissä oli käyttäjien selvä turhautuminen ohjelman toimintaan (taulukko 8). Tämä johtui osittain siitä, että koska ohjelma tuotti paljon dataa, se teki myös paljon virheitä (taulukko 8). Sisällönkuvailijat, joihin tässä yhteydessä luetaan sekä tuotantokoordinaattorit että informaatikot, eivät ole tottuneet niin suureen virheiden korjaamisen määrään. Vaikka informaatikkojen työkuva sisältää tuotantokoordinaattorien tekemien sisällönkuvailujen korjaamista, on ihmisen tekemässä sisällönkuvailussa vähemmän virheitä. Tuotantokoordinaattori arvioi tekevänsä nopeammin ja helpommin sisällönkuvailun alusta alkaen itse. Automaattisen sisällönkuvailun ohjelma teki niin paljon virheitä, että virheiden korjaamiseen olisi mennyt enemmän aikaa. Tuotantokoordinaattorin arvion mukaan koneen tuottama sisällönkuvailun onnistumisprosentti oli noin 10 prosentin luokkaa, mikä oli merkittävästi huonompi kuin mitä hän oli olettanut. Automaattisen sisällönkuvailun testauksesta pidettiin testipöytäkirjaa, johon testausta tekevät henkilöt kävivät kirjoittamassa palautetta ohjelmisto-

kehittäjille. Testipöytäkirjasta selviää, että 25.10.2016 testausta tehnyt henkilö oli kuvailut automaattisen sisällönkuvailun ohjelmalla A-studion. Testipöytäkirjan osuus, johon kirjattiin havaintoja siitä, millaista dataa automaattisen sisällönkuvailun ohjelma oli tuottanut kuvantunnistuksesta, kertoo seuraavaa:

”Studiokeskusteluosuudessa tunnistettu Tv-studio, sekä solmio ja yleismääre tukka. Alkutunnuksen grafiikassa tunnistettu koralliriutta, sametti, puuhka, sateenvarjo, vaatekauppa. Studio-osuuksissa toistuu aina pöytätennispallo”

Suurin osa tunnistetuista objekteista oli kuvailtu virheellisesti tai kuvaukset eivät olleet relevantteja. Tuotantokoordinaattori oli keskustellut automaattisen sisällönkuvailun ohjelman tuottamasta sisällönkuvailusta silloisen arkistopäällikön kanssa. Hänen mukaansa ohjelmasta ei ole hyötyä, jos vastaavuusprosentti on noinkin huono. Tuotantokoordinaattori kommentoi automaattisen sisällönkuvailun ohjelman korjaamista seuraavasti:

”Tolla ajalla kun mun piti klikata tonne ja käydä tossa ja plaa plaa plaa... ni siinä ajassa mä olisin kirjoittanut sen jo”

Virheiden korjaaminen koettiin siis hitaaksi ja turhauttavaksi työksi. Kehityspäällikkökin toteaa haastattelussaan, että virheiden korjaamiseen turhautuu nopeasti. On myös kyseenalaista, oliko virheiden korjaaminen tehty käytettävyytensä osalta niin, että käyttäjä olisi mahdollisimman helppoa korjata virheitä.

Käyttäjää tuntui virheiden määrän ja virheiden korjaamisen hankaluuden lisäksi myös turhauttavan se, että automaattiselle sisällönkuvailun ohjelmalle pitää opettaa sama asia useaan kertaan. Vaikka palautetta käyttäjien puolesta annetaan, palaute ei heti vaikuta automaattisen sisällönkuvailun ohjelmaan vaan ohjelma vaatii ylläpitoajoja. Käyttäjien antama palaute saattaa vaikuttaa sisällönkuvailuun vasta viikkojen päästä automaattisen sisällönkuvailun tuottamaan sisällönkuvailuun. Käyttäjän kannalta tämä saattaa kehityspäällikön mukaan aiheuttaa turhautumisen tunteita:

”Ihmisen kannalta mahdollisesti turhauttavaksi koettu viive siinä, että minähän sain sille, että tuo on väärin ja se tarjoaa niitä väärä silti”

Tärkeää käyttäjän turhautumisen minimoimiseksi olisi se, ettei käyttäjä joudu opettamaan toistuvasti samaa asiaa. Ongelmaksi tässä nousee ohjelman päivittämisen hitaus. Ohjelma olisi saattanut oppia tunnistamaan esimerkiksi henkilön jo muutamasta ensimmäisestä opettamiskerrasta, mutta koska päivittäminen tapahtuu hitaasti, käyttäjä ehtii toistamaan samaa asiaa ohjelmalle useamman kerran.

Turhautumista käyttäjissä saattaa myös aiheuttaa se, että koneen on vaikea todeta tunnistetun objektin merkittävyyttä ihmisille (taulukko 8). Asioiden merkityksellisyys eri konteksteissa on vaikea opettaa koneelle. Tuotantokoordinaattori kertoi harvinaisista ja hänen kokemuksensa mukaan kauniista maisemakuvista, jotka hän koki merkityksellisiksi sekä kulttuurihistoriallisesti että jälleenkäytön kannalta.

”Tos oli semmonen, et ne ajo aamulla Eestin läpi ja siinä oli tosi kauniita tämmösiä maisemakuvia. Semmosta sumua... Ja silloin se ehdotti vettä. Tommonen kuva, sehän on semmonen kuva mitä kuka tahansa voi tunnekuvana käyttää.”

Tuotantokoordinaattori siis koki, että kuva oli merkityksellinen, koska se kuvasi jotakin inhimillistä tunnetta (tunnekuva). Tämän kaltaisen merkitysten opettaminen automaattiselle sisällönkuvailulle koettiin mahdottomaksi. Kun käyttäjä olisi kuvaillut kuvan tunnelmaltaan harvinaiseksi tunnekuvaksi ja kone vedeksi, käyttäjän ja koneen välille muodostui semanttinen kuilu,

Hare et al. (2006) määrittelevät semanttisen kuilun olevan ”yhteisen näkemyksen puutetta visuaalisen datan tarjoaman informaation ja käyttäjän siitä tietyssä tilanteessa tekemän tulkinnan kanssa”. Yhteisen näkemyksen puute kuvaa hyvin semanttista kuilua. Automaattisen sisällönkuvailun ohjelma tekee multimediallisesta aineistosta jonkin tulkinnan, josta syntyy tietty määrä dataa. Ylen automaattisen kuvailun hankkeessa automaattisen sisällönkuvailun ohjelma tuotti esimerkiksi avainsanoja, kuten solmio tai pallo. Näille avainsanoille ei kuitenkaan syntynyt mitään semantiikkaa, vaan semantiikka jää täysin avainsanoja tulkitsevan ihmisen työksi. Ihmisen kuvailemalla aineistoja, on semantiikka aina kuvailussa mukana. Yingin ja kumppaneiden (2007) mukaan semanttisen kuilun aiheuttaa se, että ihmisillä on tapana käyttää korkeatasoisia konsepteja ja automaattinen konenäköön perustuva tekniikka luo matalan tason sisältöjä (värejä, tekstiä, muotoja jne.). Korkeatasoiset konseptit tulevat ihmisiltä ikään kuin luonnostaan ja tarvitsemme niitä, että voimme luoda merkityksellisiä kokonaisuuksia. Muuten sisällönkuvailu voi olla vain sanojen lista, jotka eivät tuota merkitystä eli semantiikkaa. Haren ja kumppaneiden mukaan on kuitenkin olemassa tilanteita, jossa yksinkertainen kuva on mahdollista kuvailla semanttisesti automaattisin menetelmin. Tämä onnistuu, jos kuvasta saadaan aikaiseksi objekteja tunnistamalla ja nimeämällä täydellinen semanttinen kuvailu. Tämän kaltaiset tilanteet kuitenkin lienevät harvinaisempia. Kuten aiemmin mainittiin, automaattisen sisällönkuvailun ohjelma pystyy tunnistamaan kuvapinnasta objekteina kaksi erillistä koi-

raa ja pallon, jolloin tiedonhakijan tiedontarve täyttyi tuossa tapauksessa. Näin koko tarvittava semanttisuus voitiin ilmaista avainsanojen avulla. Jos kuitenkin tiedonhakijan tarve on hiemankin abstraktimpi, esimerkiksi tiedonhakija haluaa videokuvaa onnellisesta koirasta, ei semanttisuutta tällä tavoin saada vangittua.

Kehityspäällikön mukaan kyse on siitä, että automaattisen sisällönkuvailun ohjelmalla ei ole tilanneharkintaa. Informaatikko piti keskeisenä ongelmana, ettei kone pysty luomaan yläkäsitteistöä. Voidaan olettaa, että kaikki haastatellut puhuvat samasta asiasta käyttäen hieman erilaisia termejä kuvaamaan ongelmaa. Informaatikko toteaa, että kone saattaisi pitää merkityksellisenä ajallisesti, pitkään kestävästä kuvasta traktorista. Tämä kuva olisi kuitenkin tylsää katsottavaa ihmiselle. Voidaankin kysyä, miksi pitkään jatkuva maisemakuva on ihmiselle merkityksellistä, mutta pitkään jatkuva kuva traktorista on tylsää katsottavaa? Kuvan merkitystasoa voi olla vaikea määritellä sisällönkuvailun ohjelmille siksi, että merkitsevyden ymmärtäminen on vaikeaa myös ihmiselle itselle.

Taulukko 8. Automaattisen sisällönkuvailun oppimiskyky on rajallinen

Kone tekee paljon virheitä ja oppii hitaasti, joten käyttäjä turhautuu	Koska automaattinen sisällönkuvailu tuottaa paljon dataa, on siinä paljon korjattavaa
	Koneen virheiden korjaaminen on hidasta ja turhauttavaa työtä
	Sisällönkuvailijana en halua opettaa montaa kertaa samaa asiaa
Tunnistetun asian merkittävyyden toteaminen on vaikeaa koneelle	Se milloin ja miksi jokin asia on merkittävä, on vaikea opettaa koneelle
	Koneen on vaikea arvioida tunnistamansa asian merkittävyyttä ihmisille
Automaattisen sisällönkuvailun oppiminen on sidottu ihmisiin	Automaattinen sisällönkuvailu vaatii jatkuvaa opettamista

	Koneen opettaminen vaatii paljon työtä
	Valmista sovellusta automaattiseen sisäl- lönkuvailuun ei vielä ole

7 IHMISEN JA KONEEN YHTEISTYÖ

Tässä luvussa keskitytään automaattisen sisällönkuvailun puutteisiin ja pohditaan, miten automaattista sisällönkuvailua voisi silti hyödyntää työkaluna sisällönkuvailussa.

7.1 Kone ei vielä korvaa ihmistä

Vaikka aiemmin esitetyt empiiriset havainnot Ylen hankkeesta viittaavat siihen suuntaan ettei automaattinen sisällönkuvailu tuota sisällönkuvailua riittävällä tasolla, tämä ei kuitenkaan merkitse, etteikö automaattisesta sisällönkuvailusta olisi hyötyä. Kone olisi hyvä tuottamaan lisäarvoa tiedonhakuun. Kone pystyy tuottamaan tiedonlouhintaan raakaa dataa ihmistä paremmin (taulukko 9). Kehityspäällikön mukaan automaattista sisällönkuvailua voitaisiin valjastaa hakuindeksien rikastamiseen. Informaatikko oli kehityspäällikön kanssa samoilla linjoilla pohtiessaan, miten esimerkiksi julkaisutoiminnan kehittyminen voisi hyödyntää rikasta metadataa, jota kone tuottaa.

Metadatalla tässä tutkimuksessa tarkoitetaan tietoa tiedosta. Metadata kuvaa kuvattavan asian moninaiset ominaisuudet, jotka ovat hyödyllisiä aineistojen käyttämisen, haun tai hallinnan kannalta. (Chowdhury & Chowdhury 2007, 140).

Kehityspäällikön mukaan jo osittainkin oikea metadata voisi hyödyttää tiedonhakutilanteissa. Tällöin koneelta ei vaadittaisi täydellisen semanttisen sisällönkuvailun tuottamista.

”Jos halutaan tarjota ikään kuin vaan tarttumapintaa erilaisille hakukoneille ni se voi olla ihan hyvä se semmonen iso massa jotain tietoa”

Tällöin metadata voisi auttaa ihmistä, mutta ihmissilmän näkymättömissä. Kehityspäällikön mukaan metadataa voisi piilottaa ”kulisseihin”, eikä näin ollen sen tarvitsisi olla edes ihmisluettavaa. Myös Chowdhury ja Chowdhury (2007, 151) kirjoittavat, että vaikka metadata on tärkeää aineistojen etsimisessä, loppukäyttäjät eivät yleensä näe, tai heidän ei ole tarpeen nähdä, sen aineiston metadataa, jota he etsivät.

Metadatan muokkaaminen sellaiseksi, että se suoraan olisi ihmisen luettavissa tai niin, että metadata sovitettaisiin suoraan ihmisen tekemään työhön, on huomattavan työlästä.

Ihmiselle voisi kehityspäällikön mukaan näyttää jonkinlaisen karsitun joukon metadatasta. Rikkaampi ja runsaampi kuvailu olisi lähtökohtaisesti piilotettuna. Näin metatiedot saataisiin upotettua koneellisiin prosesseihin. Kaiken kaikkiaan olisi syytä miettiä, mitä ihmiselle kannattaa ja tarvitsee automaattisen sisällönkuvailun ohjelman tuottamasta metadatasta ylipäättään näyttää.

Tiedonhaku kuitenkin hyötyisi yksityiskohtaisemmista metatiedoista (taulukko 9). Automaattisen sisällönkuvailun ohjelma tunnistisi esimerkiksi erittäin hyvin kuvassa olevat solmiot. Tämä voisi informaation mukaan hyödyttää joissakin hauissa. Informaatikko toteaa myös, että Yle Areenan tarpeisiin koneen tuottama sisällönkuvailu voi olla riittävää, ellei jopa sopivaa. Tämä kuitenkin vaatisi tiedonhakuprosessin jatkokehittelyä, koska informaation mukaan se saattaisi vaatia tiedonhakuprosesimäistä tietojen yhdistämistä.

Automaattisen sisällönkuvailun ohjelmalta vaaditaan enemmän kuin mihin se toistaiseksi pystyy (taulukko 9). Tämä saattaa johtua siitä, että ihmisen tekemään sisällönkuvailuun tottuneella henkilöllä on liian kovat vaatimukset automaattisen sisällönkuvailun ohjelman tuottamalle sisällönkuvailulle. Informaatikko arveli, että

”Se on vaan pakko hyväksyä, että se kone tuottaa 20-40% roskaa ja sitä on ihmisen tuottamaan sisällönkuvailuun tottuneen hirmu vaikea hyväksyä”

Informaatikko kertoo toivoneensa ennen testin alkua, että automaattisen sisällönkuvailun ohjelma tuottaisi laadukasta sisällönkuvailua ja että se tunnistaisi objekteja paremmin kuvapinnasta. Kehityspäällikön mukaan on väistämätöntä, että kone tekee vääriä tulkintoja ja että automaattisen sisällönkuvailun ohjelman sisällönkuvailussa on ”kuonaa” seassa. Sekä informaatikon, että tuotantokoordinaattorin haastatteluissa kävi useaan kertaan ilmi se, että koneen uskottiin tuottavan laadullisesti parempaa sisällönkuvailua ja että koneen tuottama sisällönkuvailu koettiin huonoksi, osin jopa käyttökelvottomaksi. Tämän vastanoton vuoksi voisi olla automaattiselle sisällönkuvailulle edullisempaa hyödyttää tiedonhakua ilman, että sen tuottama metadata olisi näkyvillä sisällönkuvailijalle tai edes tiedonhakijalle.

Ihmisillä on siis hyvin erilaiset odotukset siitä, millaista sisällönkuvailua koneen tulisi tehdä (taulukko 9). Kehityspäälliköllä oli haastattelun perusteella realistisempi kuva jo ennen hankkeen alkua siitä, minkälaiseen sisällönkuvailuun automaattisen sisällönkuvailun ohjelma pystyy. Tuotantokoordinaattori piti automaattisen sisällönkuvailun ohjelmaa itsensä kannalta täysin hyödyttömänä. Informaatikkokin totesi, että se sisällönkuvailu,

mitä automaattisen sisällönkuvailun ohjelma tässä kehitysvaiheessa tuottaa ei ole mitenkään riittävää. Sisällönkuvailu on edelleen tehtävä entisin keinoin manuaalisesti. Koska edelleen Ylen Arkistossa tarvitaan suorasanaisten kokonaisista lauseista muodostuva luonnollisella kielellä tehty sisällönkuvailu. Kehityspäällikkö katsoi, että hanke oli enemmänkin taustaselvitystyyppinen projekti, jonka tarkoituksena oli selvittää eri tehtäviä, joita kone voisi tehdä, mitä kyseinen palvelu tekee ylipäätään, millä laadulla ja millä tasolla.

Kone ei kuitenkaan yhdenkään haastatellun mielestä mitenkään korvaa ihmisen työtä, mutta voisi muuttaa sitä. Ainakin kone voisi tuottaa ihmisen työhön lisäarvoa (taulukko 9). Informaatikko arveli, että

”Se ei ehkä korvaa mitään, mutta mitä se tuo lisää. Mä nään sen todennäköisempänä tulevaisuuden kuvana, koska se tuo jotain lisäarvoa siihen ohjelmaan”

Pohdintaa aiheutti myös se, miten metadatan tarve tulee tulevaisuudessa muuttumaan ja voiko tätä tarvetta enää tyydyttää yksinomaan ihmislähtöisellä sisällönkuvailulla. Tarvitaanko kenties automaattisen sisällönkuvailun ohjelmaa vastaamaan tähän uuteen metadatan tarpeeseen, jota aiheuttaa esimerkiksi uudenlainen aineistojen julkaiseminen? Uudenlainen sisällönkuvailu vaatii metadatan tarpeen uudelleenkartoitukseen (taulukko 9). Informaatikon mukaan käsitys siitä, minkälaista metadataa ylipäätään tarvitaan, vaatii suuren muutoksen. Hänen mukaansa on muistettava, että Ylen tietokanta on alun perin suunniteltu ja kehitetty Yleisradion sisäisille käyttäjille. Näiden käyttäjien metadatan tarve on ollut erilainen. Informaatikon mukaan Ylen sisäisten käyttäjien yleisin syy aineistojen hakuun on se, että tietokannasta haetaan ohjelmia kokonaan uusittaviksi tai etsitään insert-käyttöön materiaalia. Insert-käyttö tarkoittaa sitä, että etsitään osia toisista ohjelmista käytettäväksi muissa ohjelmissa.

Haastatteluista on pääteltävissä se, että vielä ei ole pystytty tarkalleen hahmottamaan sitä, minkä ihmisen työvaiheen kone voisi korvata. Informaatikko myöntää, että odotukset automaattisen sisällönkuvailun suhteen eivät ole olleet realistiset:

”Meillä on ollu ehkä väärät odotuksetkin, että se korvais sen nykyisen ihmisen tekemän sisällönkuvailun”

Kehityspäällikön mukaan tähän kysymykseen liittyy tehtävänjakoa Ylen Arkistossa yleisemminkin. Kysymykseksi nousee esimerkiksi se, minkä tyyppinen metadata olisi ko-

nelähtöistä ja mikä ihmisen tekemää. Samaa pohdintaa on löydettävissä myös informaatikon haastattelusta, jossa hän pohdiskeli, muuttuuko tarve ihmisen tekemälle sisällönkuvailulle, jos automaattinen sisällönkuvailu alkaa tarjota metadataa osaksi sisällönkuvailua. Tulevaisuudessa voisi esimerkiksi olla mahdollista, että ihmisen tehtäväksi sisällönkuvailussa jää vain koneen tuottaman sisällönkuvailun validoiminen.

Taulukko 9. Kone ei vielä korvaa ihmistä, mutta mahdollistaa uudenlaisen tekemisen tavan

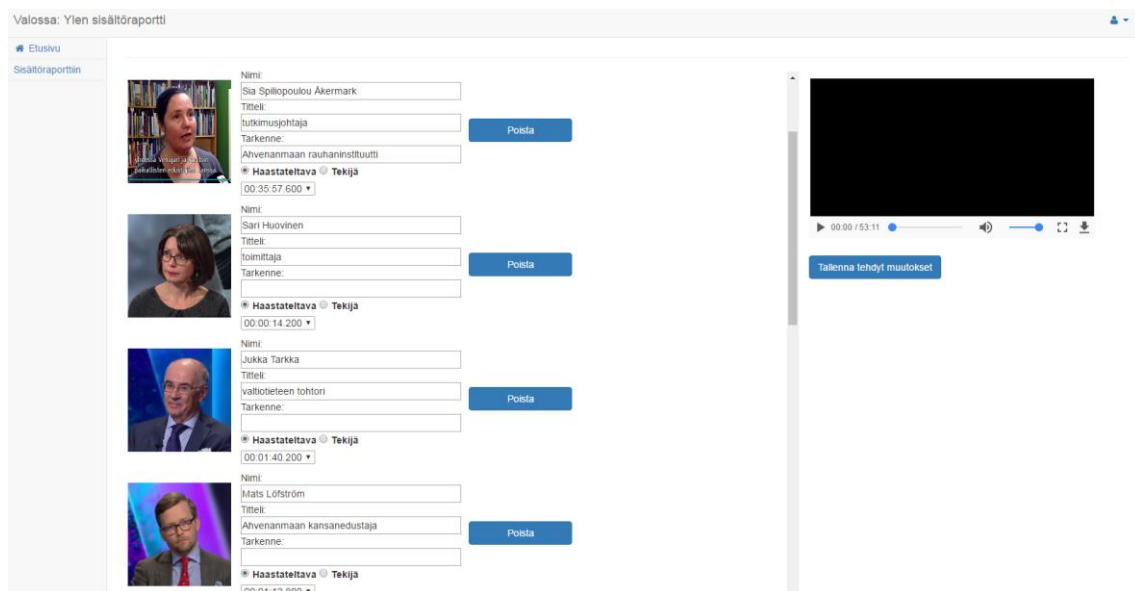
Kone on hyvä tuottamaan lisäarvoa tiedonhakuun näkymättömissä	Kone pystyy tuottamaan tiedonlouhintaan raakaa dataa ihmistä paremmin
	Metadata voisi auttaa tiedonhakua ihmisen silmän näkymättömissä
	Haku hyötyy yksityiskohtaisista metatiedoista
Automaattiselta sisällönkuvailun ohjelmalta odotetaan enemmän kuin mihin se pystyy	Ihmisen tekemään kuvailuun tottuneella on liian kovat vaatimukset koneelle
	Ihmisillä on erilaiset odotukset siitä millaista sisällönkuvailua koneen tulisi tehdä
Kone ei korvaa ihmisen työtä, mutta voisi muuttaa sitä	Kone ei korvaa ihmistä, mutta tuottaa ihmisen työhön lisäarvoa
	Uudenlainen sisällönkuvailu vaatii tarpeen (metadatan) uudelleen kartoittamisen
	Vielä ei ole pystytty hahmottamaan, minkä ihmisen työvaiheen kone voisi korvata

7.2 Automaattinen sisällönkuvailu voisi tasalaatuistaa sisällönkuvailuja

Videomateriaalissa esiintyvien henkilöiden tunnistaminen koetaan Ylessä erityisen tärkeäksi osaksi sisällönkuvailua. Koska henkilöhaut ovat yksi tavallisimpia hakuja, joita Ylen Arkiston tietokantoihin kohdistetaan. Haastatteluissa ilmeni, että vaikka muuten sisällönkuvailu saatetaan joutua tekemään summittaisesti, esimerkiksi aikaresurssin vähyden vuoksi, ohjelmassa esiintyvät henkilöiden nimet pyritään silti aina kirjoittamaan sisällönkuvailuun. Henkilöiden tunnistamisessa automaattinen sisällönkuvailu koettiin hyödylliseksi, kuten kehityspäällikkö totesi:

”Siinä se oli erittäin hyvä jotenkin se henkilöiden ympärillä kuvien ja niinku henkilöiden esiintymisen ympärillä oleva asia”

Automaattisen sisällönkuvailun ohjelma pystyi siis tuottamaan hyötyä ihmiselle siten, että ohjelma osasi määrittää koska tietty henkilö puhuu ohjelmassa ja liittää aikakoodiin hänen nimeensä. Ohjelma tuotti myös henkilöiden yhteenvetosivun, johon poimittiin kaikki ohjelmassa esiintyneet henkilöt. Kuvassa yhdeksän on nähtävillä automaattisen sisällönkuvailun ohjelman henkilöt -yhteenvetosivu.



Kuva 9. Automaattisen sisällönkuvailun ohjelman henkilöt -yhteenvetosivu.

Kuvassa yhdeksän esitetty henkilöt -yhteenvetosivu sai positiivista palautetta ja se koettiin hyödylliseksi, koska se esitti tärkeäksi koetut henkilöiden nimet ja kuvat selkeänä

kokonaisuutena. Ongelmaksi kuitenkin nousi se, ettei automaattisen sisällönkuvailun ohjelma osannut vielä yhdistää opittuja henkilöitä eri ohjelmien välillä. Tämä ominaisuus olisi ollut seuraavaksi kehitettävien ominaisuuksien listalla, mutta hankkeen lyhyden vuoksi ominaisuus ei ehtinyt mukaan tähän testiin.

Henkilöiden tunnistamisen ja oikein nimeäminen on niin tärkeä ominaisuus, että siihen toivottiin koneelta myös muunlaista apua. Informaatikko piti hyvänä, jos tekijätiedot valittiin määrämuotoisesta valikosta. Tällöin kirjoitusvirheiden määrä vähenisi (taulukko 10). Ylen automaattisen sisällönkuvailun hankkeessa ollut sisällönkuvailun ohjelma pyrki tunnistamaan henkilöiden nimiä ja titteleitä kuvassa näkyvistä planssiteksteistä. Planssitekstien tunnistaminen osoittautui kuitenkin vaativaksi tehtäväksi ja informaatikon mukaan automaattisen sisällönkuvailun ohjelma teki planssitekstintunnistuksessa huomattavan paljon virheitä. Ongelma planssitekstien tunnistamisessa johtui siitä, että se olisi pitänyt optimoida jokaiselle ohjelmatyypille erikseen, koska jokaisessa ohjelmassa on omanlaisensa graafinen ilme. Informaatikko kertoi esimerkin A-studio -ohjelmasta, jossa on a-kirjaimen tyyppinen merkki, joka kuvaa nimen eteen laitettuna sitä, että kyseessä on toimittaja. Kaikki tämän tyyppiset erikoisuudet pitäisi opettaa automaattisen sisällönkuvailun ohjelmalle erikseen. Informaatikko kuitenkin myönsi, että joidenkin henkilöiden osalta automaattisen sisällönkuvailun ohjelma helpottaisi jo nyt henkilöiden tunnistamista.

Inhimillisistä syistä johtuen ihmisen tuottaman sisällönkuvailun laatu vaihtelee (taulukko 10). Kehityspäällikön mukaan ihmisen tekemään sisällönkuvailuun vaikuttaa esimerkiksi taustat sekä se, minkälaisessa tilanteessa ihminen sisällönkuvailua tekee. Tämä johtaa siihen, ettei ihmisten tekemä sisällönkuvailu ole vertailukelpoista keskenään. Tuotantokoordinaattorilla oli kuvailussa apunaan henkilö, joka oli ollut mukana tekemässä ohjelmaa. Ohjelman tekemisessä mukana ollut henkilö pystyi antamaan kuvailuun tietoa, jota tuotantokoordinaattorilla ei ollut. Näin he pystyivät tuottamaan parempaa sisällönkuvailua ohjelmasta. Ihmisillä on eri määrä tietoa käytettävissään, kun sisällönkuvailua tehdään. Kehityspäällikön mukaan on otettava huomioon, että ihmisen tekemä sisällönkuvailu on aika subjektiivista. Samankaan ihmisen tekemä sisällönkuvailu ei ole vertailukelpoista eri päivinä, vaikka helposti syntyy sellainen ajatusharha, että ihmisen tuottama sisällönkuvailu olisi aina yhdenmukaista.

Ihmisen tekemään sisällönkuvailuun vaikuttaa tunnetila tai väsymys, mutta kone ei väsy asioiden tunnistamiseen tai sisältöjen kuvailuun. Peter Enser (2008, 534) huomauttaa, että samallekin ihmiselle tietty kuva voi tarkoittaa eri asiaa eri aikoina tai eri yhteyksissä. Enserin näkemys on tulkittavissa niin, että ihmisen tekemään kuvailuun vaikuttavat kuvailijan senhetkinen tilanne esimerkiksi kiire tai vaikkapa kuvailijan tunnetila.

Ihmiset siis väsyvät tunnistamiseen, joten informaation mukaan ei ole järkevää, että ihminen käy klikkaamassa aikakoodin paikalleen joka kerta, kun tietty henkilö näkyy kuvassa. Automaattisen sisällönkuvailun ohjelmalle tämä ei olisi ongelma. Myös aineiston käyminen useaan kertaan läpi eri näkökulmasta tai erilaisella orientaatiolla aineistoa kohtaan voisi tuottaa kehityspäällikön mukaan mielenkiintoista metadataa aineistosta. Tämä taas olisi hänen mukaansa tyypillisesti sellainen tehtävä, joka olisi ihmisen hyvin työläs hoitaa, sillä toisin kuin kone, ihminen väsyä aineistojen läpikäymiseen. Toisaalta kehityspäällikkö huomautti, että automaattisen sisällönkuvailun ohjelma saattaa tehdä liikaaakin töitä jonkin vähäpätöisemmäksi koetun asian parissa tuottaen siitä valtavasti metadataa.

Asian opittuaan kone pystyy tuottamaan aina saman kuvailun. Tässä automaattisen sisällönkuvailun testissä koneen oppiminen oli kuitenkin kiinni ihmisestä, koska automaattisen sisällönkuvailun ohjelma tunnsti vain sen, mitä oli opetettu. Informaatikko kertoi koneen pystyvän tunnistamaan tarkasti yllättäviäkin asioita, esimerkiksi lippuja ja suojatiet. Myös tuotantokoordinaattori katsoi, että kone voi oppia tunnistamaan esimerkiksi kadulla olevia ihmisiä. Varsinkin uniikit kohteet koettiin helpoksi opettaa koneelle (taulukko 10). Unikkejä kohteita ovat tässä tapauksessa esimerkiksi aikaisemminkin mainitut henkilöt sekä erityiset rakennukset. Informaatikko arveli että

”mut et se voi oppia tunnistamaan Turun Tuomiokirkon, Helsingin Uspenskin katedraalin ja tällaiset hyvin uniikit kohteet”

Kerran siis uniikin kohteen opittuaan automaattisen sisällönkuvailun ohjelma tunnistaa kohteen jokaisella kerralla. Tämä osaltaan vaikuttaa siihen, että automaattisen sisällönkuvailun ohjelma tekee aina tasalaatuisia sisällönkuvailua (taulukko 10). Kehityspäällikön sanoin:

”Ehkä siin on semmonen geneerisyys siinä koneen tekemässä. Että se on systemaattisesti jonkunlaista, on se sitten hyvää tai huonoa, oikein tai väärin.”

Tasalaatuisuus tässä yhteydessä ei siis tarkoita tasalaatuisesti hyvää. Ilman ihmisen suorittamaa korjaamista tai koneen opettamista automaattinen sisällönkuvailun ohjelma tunnistaa tasaisesti väärin objekteja ja tekee näin tasaisen laadutonta sisällönkuvailua.

Oikein kohdennettuna automaattisen sisällönkuvailun ohjelma voisi kuitenkin auttaa sisällönkuvailussa. Koneen uskottiin auttavan esimerkiksi virheiden korjauksessa. Kirjoitusvirheiden osalta ainakin oikoluku olisi tärkeä ominaisuus. Olisikin hyvä tutkia yksityiskohtaisemmin, miten ihmisen ja koneen yhteistyö saataisiin toimimaan. Tätä kehityspäällikkö toivoi todeten että

”Mitä ei oo kauheesti tutkittu tai mitä ei oo näkyny, et mikä olis se ihmisen ja koneen interaktion tapa, se rinnakkain toiminnan tapa siinä sisällönkuvailussa”

Yhteistyötä voitaisiin kehittää esimerkiksi siihen suuntaan, että automaattisen sisällönkuvailun ohjelma tuottaa dataa, jota ihminen sitten korjaa. Automaattinen sisällönkuvailu korvaisi nykyprosessissa tuotantokoordinaattorin työn sisällönkuvailun osalta. Informaattikkokin pohti, tarvitaanko ihmisen kuvailua kenties vähemmän tulevaisuudessa, jos automaattisen sisällönkuvailun ohjelma tuottaa paljon metadataa? Vaikka vielä emme ole siinä pisteessä, että ihmisen ja koneen yhteistyöstä syntyisi riittävällä tasolla olevaa sisällönkuvailua, kehityspäällikön mukaan tässä testissä on otettu ensi askeleita siihen suuntaan.

Taulukko 10. Automaattinen sisällönkuvailu voisi tasalaatuistaa sisällönkuvailuja

Henkilöiden tunnistus on erityisen tärkeää ja kone voisi jo nyt auttaa siinä	Sisällönkuvailijana tahtoisin, että kone kerran opittuaan tunnistaisi aina henkilön automaattisesti
	Erityisen tärkeää sisällönkuvailussa on tunnistaa ohjelmassa olevat henkilöt
Ihmisen tuottaman sisällönkuvailun laatu vaihtelee	Ihmiset eivät pysty tuottamaan tasalaatua sisällönkuvailua
	Toisin kuin ihminen kone ei väsy asioiden tunnistamiseen

Oikein kohdennettuna kone voisi auttaa sisällönkuvailijaa	Haluaisin, että kone korjaa virheitäni
	Olisi hyvä tutkia miten ihmisen ja koneen yhteistyö saataisiin toimimaan
Asian opittuaan kone pystyy tuottamaan aina saman kuvailun	Kone tunnistaa sen mikä sille on opetettu
	Kone oppii uniikit kohteet helposti (rannukset, ihmiset)
	Kone tekee aina tasalaatuista sisällönkuvailua

7.3 Vielä ei olla valmiita automaattisen sisällönkuvailun tuomiin muutoksiin

Sisällönkuvailijat arvostavat nykyistä tekemisen tapaa (taulukko 11). Kaikkia kolmea haastateltavaa kehoitettiin antamaan nykyisin Ylessä käytössä olevalle sisällönkuvailun ohjelmalle (Metro), arvosana. Arvosanat annettiin asteikolla 1-5, jossa 5 on erinomainen ja 1 erittäin huono. Tuotantokoordinaattori antoi Metro -järjestelmälle arvosanan 4, kehityspäällikkö arvosanan 3,5 tai 4 ja informaattikko arvosanan 3-. Metro -järjestelmää on vuosien saatossa kehitetty toimivammaksi Ylen sisällönkuvailun tarpeita varten. Informaattikko totesikin Metron olevan ”ihan hyvä työjuhta tähän nykyiseen prosessiin”. Metro -järjestelmää arvioitaessa on otettava huomioon, että käyttäjät ovat käyttäneet samaa järjestelmää pitkän aikaa ja ovat tottuneet siihen. Tuotantokoordinaattorin sanoin

”Mä tykkään niistä hakusysteemeistä ja mä oon käyttänyt sitä niin paljon, että mun mielestä se ei oo niinku hankala”

Tästä kommentista on tulkittavissa, että tuotantokoordinaattori tietää haun ehkä olevan hankala henkilölle, joka ei ole järjestelmää käyttänyt. Tämä kertoo siitä, että tuotantokoordinaattori arvelee käytettävyydessä olevan parantamisen varaa, mutta hän on tottunut järjestelmään ja on siihen näin ollen myös tyytyväinen. Informaattikko ja tuotantokoordinaattori kertovat käyttäneensä Metroa vuodesta 2009 lähtien, jolloin Metro otettiin käyttöön Ylen Arkistossa.

Totutusta luopuminen saattaa aiheuttaa muutosvastarintaa, jota sitten uudentyyppinen sisällönkuvailu joutuu kohtaamaan (taulukko 11). Kehityspäällikkö ei pitänyt vielä aikaa täysin kypsänä työkalujen suorille muutoksille. Testissä ollut automaattisen sisällönkuvailun ohjelma vertautuu väkisin vanhaan järjestelmään. Esimerkiksi kysyttäessä automaattisen sisällönkuvailun ohjelman eräästä ominaisuudesta tuotantokoordinaattori vastasi suoraan sen olevan huonompi kuin Metron. Samasta ominaisuudesta informaattikolta kysyttäessä hän vastasi näin:

”Se on hirvu erilaista siinä, kun tässä meidän nykyisessä järjestelmässä ja mä koin sen aluksi vähän outona ja ehkä jopa hitaampana. Mutta nyt kun mä sitä useamman kerran käytin, ni mä aloin hoksata, että kuinka sitä kannattaa käyttää, niin mä en usko, että se loppuviimein ois sen hitaampi ku se nykyinen, kun se on vaan erilainen”

Voidaan siis päätellä, että uuden järjestelmän arviointiin ainakin jossain määrin vaikutti se, että sisällönkuvailijat olivat tottuneet erilaiseen järjestelmään.

Nykyisistä toimintamalleista siis tämän tutkimuksen mukaan pidetään ja ongelmaksi nouseekin se, etteivät ne istu automaattisen sisällönkuvailun prosessiin (taulukko 11). Automaattisen sisällönkuvailun ohjelmaa testattiin Ylessä niin, että sillä yritettiin tehdä sisällönkuvailua aivan kuten Metro-järjestelmälläkin. Tämä johti käyttäjien turhautumiseen sillä he kokivat, etteivät voi tehdä työtään automaattisen sisällönkuvailun ohjelmalla. Haastateltujen mielestä automaattisen sisällönkuvailun ohjelma ei toimi riittävän hyvin. Automaattisen sisällönkuvailun ohjelmalle pyydettiin haastattelussa arvosana, samoin kuten Metro-järjestelmällekin. Informaatikko antoi arvosanan 1,5, tuotantokoordinaattori arvosanan 1 ja kehityspäällikkö arvosanan 2,5. Kehityspäällikön mukaan testattavana ollut automaattisen sisällönkuvailun ohjelma ei kuitenkaan ole vertailukelpoinen Metro-järjestelmän kanssa (taulukko 11). Automaattisen sisällönkuvailun ohjelmaa käyttäessä käyttäjä joutuu ottamaan kantaa asioihin, jotka eivät ole relevantteja Metro-järjestelmää käyttäessä. Esimerkkinä uusista asioista, joihin käyttäjän piti ottaa kantaa, oli koneen tekemän sisällönkuvailun korjaaminen ja sitä kautta tapahtuva koneen opettaminen. Sisällönkuvailun ohjelmien tekeminen vaatii loppukäyttäjän tuntemista. Tuotantokoordinaattorin mukaan

”Välillä ku me otetaan käyttöön uusia systeemejä, ni ne on nää insinöörit jotka niitä on kehittänyt ja sitku ne yrittää kouluttaa, ni siitä ei oikein tuu mitään”

Tuotantokoordinaattori kertoi myös huomanneensa, ettei automaattisen sisällönkuvailun ohjelmaa ole ollut tekemässä tuotantokoordinaattoria. Tällöin ohjelman käytettävyys ei ole ehkä optimoitu parhaalla mahdollisella tavalla ainakaan tuotantokoordinaattorin käyttötarpeisiin. Automaattisen sisällönkuvailun ohjelma tulisi kehittää tiiviissä yhteistyössä loppukäyttäjän kanssa (taulukko 11). Toisaalta Ylen automaattisen sisällönkuvailun testissä kyse nimenomaan oli loppukäyttäjien osallistamisesta ohjelman testaukseen. Ongelma kuitenkin todennäköisesti oli siinä, että ohjelmaa ei alusta asti aloitettu kehittämään yhteistyössä käyttäjien kanssa, vaan käyttäjät osallistuivat ohjelman testaukseen vasta kun ohjelman kehittäminen oli edennyt jo pitkälle.

Taulukko 11. Automaattisen sisällönkuvailun ohjelman kanssa tehtävä sisällönkuvailu vaatisi sellaisia toimintamallien muutoksia, joihin ei olla valmiita

Nykyiseen sisällönkuvailun prosessiin on totuttu eikä siitä haluta luopua	Sisällönkuvailijat arvostavat nykyistä tekemisen tapaa
	Uudentyyppinen sisällönkuvailu kohtaa muutosvastarintaa
	Sisällönkuvailua on jo pitkään tehty samalla kaavalla
Nykyiset sisällönkuvailun toimintamallit eivät sovi automaattiseen sisällönkuvailuun	Sisällönkuvailun ohjelmien tekeminen vaatii loppukäyttäjän tuntemista
	Automaattisen sisällönkuvailun ohjelma ei ole vertailukelpoinen vanhan kanssa
	Yritin tehdä työni automaattisen sisällönkuvailun kanssa mutta en pystynyt
	Uudentyyppinen sisällönkuvailu vaatisi työn uudelleen organisoinnin

8 YHTEENVETO

Tässä luvussa esitellään tutkimuksen johtopäätökset sekä käydään läpi millaista jatkotutkimusta voitaisiin tehdä.

8.1 Johtopäätökset

Tässä tutkimuksessa käytiin läpi miten Ylen sisällönkuvailun parissa työskentelevät henkilöt kokivat automaattisen sisällönkuvailun testauksen. Tutkimuskysymykset olivat, (1) miten nykyinen multimediallisten aineistojen sisällönkuvailu Ylessä toimii tällä hetkellä, (2) miten automaattinen sisällönkuvailun ohjelma toimi Ylen videoaineistojen sisällönkuvailussa, (3) mitä automaattisen sisällönkuvailun ohjelman tuottamasta sisällönkuvailusta puuttuu ja (4) miten automaattista sisällönkuvailua voitaisiin hyödyntää helpottamaan sisällönkuvailijoiden työtä.

Ylessä multimediallisten aineistojen sisällönkuvailu on vuosien saatossa hioutunut prosessi. Sisällönkuvailua sekä ohjelmaa, jolla sisällönkuvailua tehdään, on kehitetty jatkuvasti tehokkaampaan ja laadukkaampaan suuntaan. Tuotantokoordinaattorit ja informaatikot pystyvät yhteistyössä tekemään semanttisesti rikasta kuvailua, joka riittää ainakin Ylen sisäiseen käyttöön. Nykyisessä sisällönkuvailun prosessissa on kuitenkin ongelmia. Suurin niistä on resurssipula verrattuna aineistojen suureen määrään. Ihmisvoimin tehty sisällönkuvailu ei ole kustannustehokasta ja koska aineisto määrä on valtava, ei kaikkia aineistoja saada kuvailtua riittävällä tarkkuudella. Juuri näihin ongelmiin haettiin apua automaattisen sisällönkuvailun ohjelmalta.

Tutkimuksessa kuitenkin kävi ilmi, ettei automaattinen sisällönkuvailu ainakaan tässä testissä toiminut sillä tasolla, että siitä olisi ollut merkittävää hyötyä sisällönkuvailijoille. Automaattisen sisällönkuvailun ohjelma tuotti sisällönkuvailua, joka suurelta osin oli virheellistä. Sisällönkuvailijoilla meni tutkimuksen mukaan enemmän aikaa ohjelman tuottaman sisällönkuvailun korjaamiseen kuin jos se olisi tehty alusta alkaen itse. Automaattisen sisällönkuvailun ohjelman opettaminen koettiin raskaaksi, eikä ollut selvyyttä kuinka merkittävästä prosessista olisi ollut kyse, jotta automaattinen sisällönkuvailu olisi saatu toimimaan riittävän hyvin.

Automaattisen sisällönkuvailun ohjelma kykeni opittuaan tunnistamaan kohtalaisen hyvin ihmisen kasvot ja liittämään nimen ja kasvot yhteen. Tämä koettiin tutkimuksessa automaattisen sisällönkuvailun ohjelman parhaaksi ominaisuudeksi. Ohjelmassa kuitenkin esiintyi merkittäviä puutteita. Ylen sisällönkuvailijat ovat työssään tottuneet käsittelemään semanttisesti rikasta ihmisen tekemää sisällönkuvailua, eikä automaattisen sisällönkuvailun ohjelma pystynyt tähän lainkaan. Ohjelmaa ei voitu ottaa käyttöön, koska sisällönkuvailun semantiikasta ei Ylellä olla valmiita luopumaan.

Automaattisen sisällönkuvailun ohjelman tuottaman metadatan voidaan kuitenkin olettaa jossain määrin hyödyttävän tiedonhakijaa. Ohjelma pystyy käymään aineiston useaan kertaan läpi ja kuvailemaan aineiston usean eri näkökulman avulla. Tämän kaltaista toimintaa ei voida olettaa ihmisen tekevän. Automaattinen sisällönkuvailu voisi siis luoda metadatta, jota hyödynnettäisiin sitten niin sisällönkuvailijan kuin tiedonhakijan näkömäämissä.

Tärkein löydös tässä tutkimuksessa oli se, ettei automaattinen sisällönkuvailu vielä ole kykenevä korvaamaan ihmisen tekemää työtä sisällönkuvailussa Ylellä. Prosesseja muuttamalla ja metadatan tarvetta uudelleen miettimällä automaattinen sisällönkuvailu voisi tuoda ihmisen tekemään työhön jonkinlaista lisäarvoa.

Automaattisen sisällönkuvailun hanke oli Ylessä testajamäärältään suhteellisen pieni ja testauksessa ollut automaattisen sisällönkuvailun ohjelma oli vielä varhaisessa kehitysvaiheessa oleva versio eikä valmis tuote. Tästä johtuen myös tässä tutkimuksessa tutkimusotos jäi pieneksi. Tärkeintä tutkimuksen luotettavuuden kannalta oli se, että tutkimukseen haastateltiin kolmea sisällönkuvailun kannalta tärkeää työroolia. Näin saatiin otos siitä miten automaattinen sisällönkuvailu koettiin yli työroolirajojen.

Tutkimustulokset eivät välttämättä ole laajasti yleistettävissä. Ne antavat kuitenkin oikeellista viitettä sekä siitä millaisella tasolla automaattinen sisällönkuvailu on että siitä, miten automaattiseen sisällönkuvailuun suhtautuvat sisällönkuvailun ammattilaiset. Henkilöt, joita tässä tutkimuksessa haastateltiin, olivat kaikki pitkään sisällönkuvailun parissa työskennelleitä alan ammattilaisia. Tästä syystä heidän kokemuksensa ja ajatuksensa automaattisesta sisällönkuvailusta ovat valideja ja antavat luotettavuutta koko tutkimukselle. Vaikka tutkimus siis jäi tutkimusotokseltaan pieneksi, se sisälsi harvinaisia tutkimustuloksia, jotka ovat haastateltavien ammattiosaamisen ja työroolien kautta valideja.

Automaattisen sisällönkuvailun tutkimus on verrattain vähäistä ja varsinkin käyttäjäkokemusten kartoittaminen on vielä pienimuotoista. Käyttäjäkeskeistä tutkimusta automaattisen sisällönkuvailun saralta ovat tehneet esimerkiksi Vrochidis ja kumppanit (2017). Heidän tutkimuksensa keskittyi siihen, miten käyttäjä toimii automaattisen sisällönkuvailun ohjelman kanssa. Käyttäjien reaktioita seurattiin tutkimuksessa esimerkiksi katseen seurannan avulla. Heidänkään tutkimus ei kuitenkaan ota kantaa siihen, mikä oli käyttäjien kokemus automaattisen sisällönkuvailun ohjelmasta tai miten automaattinen sisällönkuvailu vertautuu manuaalisen sisällönkuvailun ohjelman kanssa toimimiseen.

Koska aiempaa käyttäjäkeskeistä tutkimusta automaattisesta sisällönkuvailusta on hyvin vähän ovat tämän tutkielman tulokset arvokkaita. Merkittäviä tuloksia saatiin muun muassa siitä miten sisällönkuvailijat kokivat aineistojen heterogeenisyyden vaikutukset automaattiseen sisällönkuvailuun. Tärkeää oli myös sen pohtiminen, miten manuaalisen ja automaattisen sisällönkuvailun yhdistämisestä saadaan monia julkaisumuotoja palveleva kokonaisuus. Molemmista sekä automaattisesta sisällönkuvailusta että ihmisen tekemästä sisällönkuvailusta pitäisi pystyä poimimaan parhaat puolet. Siinä missä automaattinen sisällönkuvailu kykenee tuottamaan valtavan määrän metadataa ja käymään aineistoja useamman kerran läpi, ihminen kykenee tuottamaan sisällönkuvailuun sekä ylemmää tasoa että sisällönkuvailussa tarvittavaa semanttista merkitystä. Molemmat sisällönkuvailut tuovat prosessiin sellaista lisäarvoa, jota toinen ei ole kykenevä tuottamaan. Tehokkaalla koneen ja ihmisen yhteistyöllä saataisiin kapenemaan automaattisen sisällönkuvailun aiheuttama semanttinen kuilu. Semanttisen kuilun ongelmat ovat tuttuja jo aikaisemmista tutkimuksista kuten Haren ja kumppaneiden (2006) tai Enser (1999 sekä 2008). Tutkielmassa on kuitenkin uusi, eittämättä tärkeä nimenomaan käyttäjien kokemukseen pohjautuva näkökulma automaattiseen sisällönkuvailuun.

Tutkimuksessa kokeiltiin rohkeasti tutkimuksen tieteenalalle harvinaista tutkimusmenetelmää. Menetelmä auttoi nostamaan aineistosta tuoreita havaintoja nykyisestä sisällönkuvailun tuottamisen prosessista, automaattisesta sisällönkuvailusta sekä edellä mainittujen mahdollisesta yhdistämisestä. Affinity diagram todettiin tässä tutkimuksessa hyväksi metodiksi käyttäjäkokemusten kartoittamiseen, koska se ei hukkaa aineistosta mitään, vaan pitää kaiken aineiston mukana koko analyysivaiheen läpi. Affinity diagram -metodin avulla saatiin pienehköstä haastatteluutoksesta suurehko aineistomassa, joka käsiteltiin hukkaamatta käyttäjien tuomaa varianssia. Affinity diagram on siis tämän tutkielman perusteella varteenotettava metodi käytettäväksi myös jatkotutkimuksissa.

8.2 Jatkotutkimus

Tässä tutkimuksessa päästiin alkuun siinä, miten automaattista sisällönkuvailua voitaisiin hyödyntää Ylen sisällönkuvailun apuna. Jatkotutkimuksessa voitaisiinkin lähteä ratkaisemaan tätä ongelmaa syvemmin. Miten käytännössä automaattisen sisällönkuvailun tuottama data hyödyttäisi tiedonhakijaa? Miten metadatan tarpeen muuttuminen vaikuttaa sisällönkuvailuun tulevaisuudessa? Tai miten tiedonhakija kokee koneellisesti tuotetun metadatan? Tässä tutkielmassa esitettiin arvio siitä, että tiedonhaku voisi hyötyä koneen tuottamasta metadatasta tiedonhakijan näkymättömissä. Asiaa olisi kuitenkin hyvä tutkia lisää. Lisäksi erityisen kiinnostava oli ajatus siitä, että automaattisen sisällönkuvailun ohjelma voisi kuvailla aineistoja useampaan kertaan. Jatkossa olisikin hyvä selvittää, mitä lisäarvoa aineistojen kuvaileminen automaattisin menetelmin useampaan kertaan toisi.

Vaikka multimediallisten aineistojen ja varsinkin kuvan kuvailun tutkimusta on tehty jo vuosikymmeniä, ei vieläkään ole saatu täysin ratkaistua semanttisen kuvailun ongelmaa. Vastausta siihen, miten automaattinen sisällönkuvailun ohjelma saataisiin ymmärtämään kuvassa olevia merkityksiä tai objektien suhteita toisiinsa, ei ole toistaiseksi saatu. Sisällönkuvailu ei ole tämän tutkielman perusteella riittävää, jos se ei sisällä semantiikkaa. Kuitenkin semantiikan määrittelemisen automaattisen sisällönkuvailun ohjelmalle koettiin vaikeaksi. Tulevaisuudessa olisi hyvä pohtia, miten semantiikka voitaisiin määritellä niin, että automaattisen sisällönkuvailun ohjelma voisi sitä hyödyntää.

Tekoäly ja koneoppiminen ovat tieteenaloja, jotka kehittyvät jatkuvasti. On kiinnostavaa nähdä, miten esimerkiksi neuroverkot tulevaisuudessa mahdollisesti kehittävät sisällönkuvailua. Tämän tutkielman kohteena olleesta automaattisen sisällönkuvailun ohjelmasta puuttui kokonaan puheentunnistus. Olisi kiinnostavaa tietää, miten tässä yhteydessä puheentunnistus vaikuttaisi automaattisen sisällönkuvailun tuottamaan kuvailuun.

Tämä tutkimus oli luonteeltaan harvinainen ja sen tutkimustyyli oli tieteenalalleen uusi. Sen vuoksi tutkielma myös jättää paljon jatkokysymyksiä, joita olisi hyvä tutkia jatkossa.

LÄHTEET

- Beyer, H. 2010, User-Centered Agile Methods, Morgan & Claypool, San Rafael.
- Brinkmann, s. & kvale, s. 2005, "Confronting the Ethics of Qualitative Research", Journal of Constructivist Psychology, vol. 18, no. 2, pp. 157-181.
- Brunelli, R., Mich, O. & Modena, C.M. 1999, "A Survey on the Automatic Indexing of Video Data", Journal of Visual Communication and Image Representation, vol. 10, no. 2, pp. 78-112.
- Chowdhury, G.G. & Chowdhury, S. 2007, Organizing information: from the shelf to the Web, Facet, London.
- Colombo, C., Del Bimbo, A. & Pala, P. 1999, "Semantics in visual information retrieval", IEEE MultiMedia, vol. 6, no. 3, pp. 38-53.
- Dalakleidi, K., Dasiopoulou, S., Stoilos, G., Tzouvaras, V., Stamou, G. & Kompatsiaris, Y. 2011 Semantic Representation of Multimedia Content. In Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 18-49.
- Enser P. 2008 "The evolution of visual information retrieval", Journal of Information Science, vol. 34, no. 4, pp. 531-546.
- Enser P. 1999 Visual image retrieval: seeking the alliance of concept-based and content-based paradigms, Journal of Information Science, vol 26, no 4, pp. 199-210.
- Enser, P. & Sandom, C. 2002, "Retrieval of archival moving imagery - CBIR outside the frame?", IMAGE AND VIDEO RETRIEVAL, vol. 2383, pp. 206-214
- Halonen J. & Kajo H. 2015 Sisällönkuvailuohjeita - urheiluohjelmat. Sisällönkuvailu –pikaohje Yle Arkisto. Ylen sisäinen asiakirja.
- Hare J. S., Lewis P. H., Enser P. G.B & Sandom C. J. 2006 Mind the Gap: Another look at the problem of the semantic gap in image retrieval. Proceedings of SPIE - The International Society for Optical Engineering
- Hirsjärvi, S. & Hurme, H. 2008. Tutkimushaastattelu: Teemahaastattelun teoria ja käytäntö. Helsinki: Gaudeamus Helsinki University Press.
- Holtzblatt, K. & Beyer, H. 2015;2014;, Contextual design: evolved, Morgan & Claypool, San Rafael.

Holtzblatt, K., Wendell, J.B. & Wood, S. 2005;2004;., Rapid contextual design: a how-to guide to key techniques for user-centered design, Elsevier/Morgan Kaufmann, San Francisco.

Järvelin K., Sormunen E. & Mäkinen I. (toim.) 2000 Tiedon tie: Johdatus informaatiotutkimukseen. Dokumentit kateissa? Tiedon tallennus ja haku avuksi. Helsinki: BTJ Kirjastopalvelu Oy. pp. 110-142.

Järvinen, J & Walter, M. 2011, Service Design–menetelmien toimivuus moniaistisuuden näkökulmasta Case: Linnanmäki. Laurea-ammattikorkeakoulu.

Jørgensen, C., Stvilia, B. & Wu, S. 2014, Assessing the relationships among tag syntax, semantics, and perceived usefulness, Journal of the Association for Information Science and Technology, vol. 65, no. 4, pp. 836.

Kajo, H. 2018, suullinen tiedonanto 23.04.2018

Karjalainen, T. 2007, Yhdistä ideointityökalulla luovan ajattelun eri ulottuvuudet – Aivoriähi, ryhmitelykaavio sekä kalanruokaavio. <http://www.qk-karjalainen.fi/fi/artikkelit/luova-ajattelu/> (Luettu 24.04.2018)

Khan, M. & Gotoh, Y. 2017 Generating natural language tags for video information management, Machine vision and applications, vol. 28, no. 3-4, pp. 243-265.

Kvale, S. 2006, "Dominance Through Interviews and Dialogues", Qualitative Inquiry, vol. 12, no. 3, pp. 480-500.

Laki kulttuuriaineistojen tallettamisesta ja säilyttämisestä, 28.12.2007/1433. <http://www.finlex.fi/fi/laki/ajantasa/2007/20071433> (Luettu 01.04.2018)

Liu, Y., Zhang, D., Lu, G. & Ma, W. 2007, "A survey of content-based image retrieval with high-level semantics", Pattern Recognition, vol. 40, no. 1, pp. 262-282.

Markkula, M. & Sormunen, E. 2006, "Video needs at the different stages of television program making process", ACM, , pp. 111.

Naphide, H.R. & Huang, T.S. 2001, "A probabilistic framework for semantic video indexing, filtering, and retrieval", IEEE Transactions on Multimedia, vol. 3, no. 1, pp. 141-151.

Rautiainen M. 2006 Content-based search and browsing in semantic multimedia retrieval. Faculty of technology, department of electrical and information engineering, University of Oulu

Ridley, D.D., Goker, A., Davies, J. & Books24x7, I. 2009, Information Retrieval: Searching in the 21st Century, 1. Aufl.;1; edn, Wiley, New York.

Saarikoski L. 2016 Videon sisältöanalyysi PoC 2016 loppuraportti. Ylen sisäinen asiakirja.

Saarikoski L. 2016, Lauri Saarikoski: Yle valjastaa tekoälyn media-alan voimavaraksi kansainvälisessä tutkimushankkeessa. <https://yle.fi/aihe/artikkeli/2018/01/18/lauri-saarikoski-yle-valjastaa-tekoalyn-media-alan-voimavaraksi> (Luettu 31.03.2018).

Scupin, R. 1997. The KJ Method: A technique for analyzing data derived from Japanese ethnology. Article in Human organization Vol 56, No. 2

Smith J. R. 2001 Quantitative Assessment of image retrieval effectiveness. Journal of the American Society for information Science and Technology

Stickdorn, M., Lawrence, A., Hormess, M.E. & Schneider, J. 2018, This is service design doing: applying service design thinking in the real world : a practitioners' handbook, First edn, O'Reilly Media, Inc, Sebastopol, CA.

Tanni M. 2003 Digitaalisen videon automaattinen sisällönkuvailu TV-uutisissa ja videotiedonhaun käyttöliittymät. Informaatiotutkimuksen laitos Tampereen Yliopisto.

Tuomi, J. & Sarajärvi, A. 2018, Laadullinen tutkimus ja sisällönanalyysi, Uudistettu laitos. edn, Tammi, Helsinki.

Tutkimuseettinen neuvottelukunta 2009. Humanistisen, yhteiskuntatieteellisen ja käyttäytymistieteellisen tutkimuksen eettiset periaatteet ja ehdotus eettisen ennakkoarvioinnin järjestämiseksi <http://www.tenk.fi/sites/tenk.fi/files/eettisetperiaatteet.pdf> (Luettu 22.04.2018)

2016. Tällainen Yle on työnantajana <https://yle.fi/aihe/artikkeli/2016/02/22/tallinen-yle-tyonantajana> (Luettu 14.04.2018).

Vakkari, P. 1999, "Task complexity, problem structure and information actions. Integrating studies on information seeking and retrieval", Information Processing and Management, vol. 35, no. 6, pp. 819-837.

Viljanen K. & Mattila S. 2017 Tekoälyllä Ylen TV- ja radiosisällöt löydettäväksi. <http://sites.flockler.com/yle-beta/blogit/tekoalyla-ysten-tv-ja-radiosisallot-loydettavaksi> (Luettu 11.02.2018).

Virtanen P. 2016 Sisältöjen automaattinen analyysi & Yle. <http://sites.flockler.com/yle-beta/blogit/saltojen-automaattinen-analyysi-and-yle> (Luettu 11.02.2018).

Vrochidis, S., Patras, I. & Kompatsiaris, I. 2017;2016;, "Gaze movement-driven random forests for query clustering in automatic video annotation", Multimedia Tools and Applications, vol. 76, no. 2, pp. 2861-2889.

Yle (2018). Arkiston historiaa. <https://yle.fi/aihe/artikkeli/2015/02/19/arkistosta-aarteita-loyda-oh-jelma-tai-hanki-aineistoa> (Luettu 18.04.2018)

Yle (2017). Suomalaisten keskellä. view.24mags.com/mobilev/635b925edec41f121720ca00007cdcf3#/page=1 (Luettu 20.04.2018).

Zhou, X.S. & Huang, T.S. 2002, "Unifying keywords and visual contents in image retrieval", IEEE MultiMedia, vol. 9, no. 2, pp. 23-33.

Haastateltavat:

Ylen tuotantokoordinaattori 06.02.2017

Ylen Kehityspäällikkö 20.01.2017

Ylen Arkiston informaatikko 12.12.2016

LIITE 1:

Ensimmäisenä haastattelussa käydään läpi haastateltavan perustiedot, ja tämän jälkeen arvioidaan nykyisin käytössä olevaa ohjelmaa ja sen käytettävyyttä niiltä osin kuin se on oleellista tämän tutkimuksen puitteissa. Vasta viimeiseksi arvioidaan testaamisvaiheessa olevaa ohjelmaa ja sen toimivuutta siihen tarkoitukseen mihin se olisi kunkin haastateltavan kohdalla tulossa.

Haastattelu

Perustiedot

1. Ammattinimike?
2. Työnkuva niiltä osin kuin se liittyy sisällönkuvailuun.
3. Kauanko olet käyttänyt nykyisin käytössä olevaa ohjelmaa sisällönkuvailuun?

Nykyisin käytössä oleva ohjelma

1. Kuinka paljon käytät sisällönkuvailuun tarkoitettua ohjelmaa?
 - a. Kuinka kauan sinulle menee noin aikaa ohjelman (esim. A-studio), kuvailuun?
2. Mitkä ovat mielestäsi parhaat ominaisuudet nykyisin käytössä olevassa ohjelmassa?
 - a. Tarkennus, jos saisit valita nykyisin käytössä olevasta ohjelmasta ominaisuuksia uuteen ohjelmaan, mitä ne olisivat?
3. Mitkä ovat mielestäsi toimimattomimmat osuudet nykyisin käytössä olevassa ohjelmassa?
 - a. Tarkennus, jos saisit valita nykyisin käytössä olevasta ohjelmasta ominaisuuksia jotka eivät tulisi uuteen ohjelmaan, mitä ne olisivat?
4. Anna arvosana vanhalle järjestelmälle 1-5.

Nyt testattavana oleva ohjelma

1. Millaisia odotuksia sinulla oli automaattisen sisällönkuvailun osalta ennen testauksen aloittamista?
2. Kuinka paljon olet ehtinyt tehdä testauksia?
 - a. Kuinka monta ohjelmaa noin.
 - b. Kuinka kauan sinulla on mennyt aikaa kuvailuun uuden ohjelman kanssa.
 - c. Luuletko ajan nopeutuvan, kun ohjelman käyttäminen tulee tutummaksi (→ voiko käyttäminen nopeutua nopeammaksi kuin vanhalla ohjelmalla?)
3. Mitkä asiat ovat olleet erityisen positiivisia ohjelman toteutuksessa?
4. Mitkä asiat ovat olleet erityisen negatiivisia?

5. Miten ohjelman negatiiviset ja positiiviset puolet suhteutuvat sinulla olleisiin odotuksiin kuvailuohjelman suhteen?
6. Mitkä ovat sinun mielestäsi eniten kehitystä kaipaavia kohtia testattavassa ohjelmassa?
7. Anna testattavalle sovellukselle arvosana 1-5

LIITE 2:

Jälleenkäyttö on mahdollista vain jos sisällönkuvailu on tehty hyvin	Videoaineistoista oleellista on tietää mitä kuvassa näkyy	Sisällönkuvailussa pitää kertoa mitä kuvassa näkyy
	Erityyppiset tiedonhakutilanteet vaativat erityyppiset sisällönkuvailut	Yleisellä tasolla oleva tiedonhaku vaatii yleisentason kuvailun
		Hyvin tarkka tiedonhaku vaatii tarkan sisällönkuvailun
	Sisällönkuvailua tehdään jälleenkäytön vuoksi	Erityyppiset tiedonhakijat tarvitsevat erilaiset sisällönkuvailut
		Sisällönkuvailun toimivuutta arvioidaan tiedonhaun kautta
		Sisällönkuvailun tärkein motivaattori on jälleenkäytön mahdollistaminen
Videoaineistojen kuvailussa aika on otettava huomioon	Kuinka kauan ja suuresti objekti on kuvassa vaikuttaa sen relevanssiin	Objektin roolin suuruus vaikuttaa sen relevanssiin
		Kuinka kauan objekti on kuvassa vaikuttaa kohteen kuvauksen relevanssiin
		Yhden otoksen pituuden pitää vaikuttaa relevanssiin
	Tiedonhakijalle on tärkeää että kuvailu on sidottu aikaan	Tiedonhakijana haluan päästä suoraan katsomaan oikeaa hakemaani videosta
		Kuvailu pitää sitoa aikaan (aikakoodiin)
	Jotta voin sisällönkuvailla minun pitää hahmottaa ohjelman kokonaisuus ja sen osat	Koneen pitäisi pystyä medioiden avulla hahmottamaan ajallista dimensiota
		Ohjelman kokonaisuuden hahmottaminen on tärkeää
		Ohjelman osioiden tunnistaminen on tärkeää
		Ohjelman jakaminen osiin on tärkeä osa sisällönkuvailua

Automaattinen sisällönkuvailu voisi tasalaatuistaa sisällönkuvailuja	Henkilöiden tunnistus on erityisen tärkeää ja kone voisi jo nyt auttaa siinä	Sisällönkuvailijana tahtoisin että kone kerran opittuaan tunnistaisi aina henkilön automaattisesti
		Erityisen tärkeää sisällönkuvailussa on tunnistaa ohjelmassa olevat henkilöt
	Ihmisen tuottaman sisällönkuvailun laatu vaihtelee	Ihmiset eivät pysty tuottamaan tasalaatuista sisällönkuvailua
		Toisin kuin ihmiset kone ei väsy asioiden tunnistamiseen
	Oikein kohdennettuna kone voisi auttaa sisällönkuvailijaa	Haluan että kone korjaa virheitäni
		Jatkossa olisi hyvä tutkia miten ihmisen ja koneen yhteistyö saataisiin toimimaan
	Asian opittuaan kone pystyy tuottamaan aina saman kuvailun	Kone tunnistaa sen mikä sille on opetettu
		Kone oppii uniikit kohteet helposti (henkilöt, rakennukset)
		Kone tekee aina tasalaatuista sisällönkuvailua
Aineistot eivät ole homogeenisiä ja tarvitsevat siksi erityyppisiä kuvailuja	Sisältöjä pitää kuvailla monella tavalla jotta aineistojen kirjo tulee huomioitua	Joistain aineistoista menetetään jotakin merkittävää jos ne kuvaillaan liian yleisellä tasolla
		Erityyppiset videoaineistot tarvitsevat erilaista sisällönkuvailua
	Aineistojen arvo vaihtelee jälleenkäyttöarvon ja merkittävyyden vuoksi	Arvokkaiden aineistojen sisällönkuvailu on historiallisesti merkittävää
		Toiset aineistot vaikuttavat enemmän sisällönkuvailun kehitykseen

Sisällönkuvailu on mahdollista vain yhdistämällä taustatietoja ja eri medioista kerättyä dataa	Koneelle ja ihmiselle yksityiskohtainen kuvailu on vaikeaa	Ihan niin kuin ihmisenkin koneenkin pitää oppia tunnistamaan asioita
		Yksityiskohtainen sisällönkuvailu vaatii ihmiseltäkin erikoistietämystä
		Kone ei tuota riittävästi yksityiskohtaista tietoa
	Tehokas järjestelmä osaa erotella ja yhdistää tietoja	Automaattisen sisällönkuvailun pitäisi osata yhdistää monen eritietolähteen tiedot
		Tehokas sis.kuvailunjärjestelmä hyödyntää myös aiemmin tuotettua dataa
		Eri tietoja pitää pystyä erittelemään toisistaan
		Automaattisen sisällönkuvailun pitäisi osata yhdistää eri medioista saadut tiedot
	Puheentunnistus voisi auttaa automaattisessa sisällönkuvailussa	Puheentunnistus voisi auttaa relevanssin arvioinnissa
		Meillä on kokemusta siitä, että puheentunnistus helpottaa työtä
	Taustalla toimiva ontologia vaikuttaa paljon sisällönkuvailun laatuun. Siksi sen pitäisi olla näkyvä käyttäjille	Ontologian laatu vaikuttaa koneen tuottaman sisällönkuvailun laatuun
		Ontologioiden ja taksonomioiden tulee olla läpinäkyviä asiakkaalle
	Tekstitiedostot ovat oleellinen osa sisällönkuvailua	Tekstitiedoston puuttuminen vaikeuttaa sisällönkuvailua
		Tekstitiedostojen avulla kone tuottaa parempaa sisällönkuvailua

Automaattisen sisällönkuvailun ohjelman kanssa tehtävä sisällönkuvailu vaatisi sellaisia toimintamallimuutoksia joihin ei olla valmiita	Nykyiseen sisällönkuvailun prosessiin ollaan totuttu eikä siitä haluta luopua	Sisällönkuvailijat arvostavat nykyistä tekemisen tapaa
		Uuden tyyppinen sisällönkuvailu kohtaa muutosvastarintaa
		Sisällönkuvailua on jo pitkään tehty samalla kaavalla
	Nykyiset sisällönkuvailun toimintamallit eivät sovi automaattiseen sisällönkuvailuun	Sisällönkuvailun ohjelmien tekeminen vaatii loppukäyttäjän tuntemista
		Automaattinen sisällönkuvailun ohjelma ei ole vertailukelpoinen vanhan kanssa
		Automaattinen sisällönkuvailun ohjelma ei toimi vielä riittävän hyvin
		Yritin tehdä työni automaattisen sisällönkuvailun kanssa mutta en pystynyt
		Uuden tyyppinen sisällönkuvailu vaatisi työn uudelleen organisoinnin

Automaattisen sisällönkuvailun oppiskyky on rajallinen	Kone tekee paljon virheitä ja oppii hitaasti joten käyttäjä turhautuu	Koska automaatti tuottaa paljon dataa siinä on myös paljon korjattavaa
		Koneen virheiden korjaaminen on hidasta ja turhauttavaa työtä
		Sisällönkuvailijana en halua opettaa koneelle montaa kertaa samaa asiaa
	Tunnistetun asian merkittävyyden toteaminen on vaikeaa koneelle	Se milloin ja miksi jokin asia on merkittävä on vaikea opettaa koneelle
		Koneen on vaikea arvioida tunnistamansa asian merkittävyyttä ihmiselle
	Automaattisen sisällönkuvailun oppiminen on sidottu ihmisiin	Automaattinen sisällönkuvailu vaatii jatkuvaa opettamista
		Koneen opettaminen vaatii paljon työtä
		Valmista sovellusta automaattiseen sisällönkuvailuun ei vielä ole

Sisällönkuvailijana haluan tehdä mahdollisimman hyvää työtä mahdollisimman tehokkaasti	Sisällönkuvailija haluaa tehdä työnsä tehokkaasti kerralla kuntoon	Sisällönkuvailijoilta halutaan säästää aikaa
		Sisällönkuvailija ei halua tehdä uudelleen samaa sisällönkuvailua
	Sisällönkuvailussa on yhteneviä prosessinvaiheita riippumatta siitä tekeekö kuvailun kone vai ihminen	Niin ihmisenkin kuin koneenkin tuottamaa sisällönkuvailua joutuu korjaamaan
		Sekä kone, että ihminen tarvitsevat ohjeet siihen millaista kuvailua halutaan
		Sisällönkuvailu käy läpi ihmisen tekemän tarkistusprosessin tälläkin hetkellä
	Haluan hyödyntää kerran kirjattua tietoa mahdollisimman paljon	En halua syöttää montaa kertaa samoja tietoja
		Jos tietoja on jo valmiina en halua syöttää niitä uudelleen
Ihminen kertoo myös mitä kuvassa ei näy, mutta mitä se on ja mihin se liittyy (konteksti)	Ihminen tarvitaan tuottamaan kuvailuun se mitä kuvassa ei näy	Aihe antaa pohjan sisällönkuvailulle
		Sisällönkuvailija nostaa kuvailun tasoa yleisemmäksi asiasanojen avulla
		Ihminen pyrkii tuottamaan kuvailuun tarvittava yleisemmän tason
	Ilman kontekstia kuvailusta jää puuttumaan oleellista tietoa	Ihminen tuo sisällönkuvailuun kontekstin avulla semantiikkaa
		Konteksti on sisällönkuvailussa välttämätöntä